5

Quality assurance of laboratory performance

R. Wood, Food Standards Agency, UK and J. E. L. Corry, University of Bristol, UK

5.1 Introduction

It is now universally recognised as essential that a laboratory produces and reports data which are fit-for-purpose, that is fit for the intended use by the customer of the laboratory. For a laboratory to produce consistently reliable data it must implement an appropriate programme of quality assurance measures. Such measures are now required in the European Union (EU) by virtue of legislation for food control work, by the Codex Alimentarius Commission for laboratories involved in the import/export of foodstuffs and, in the case of the United Kingdom (UK) Food Standards Agency, in the requirements for contractors undertaking survey work. Thus customers now demand of providers of analytical data that their data meet established quality requirements. These are described below. The significance of the measures identified are then described and the future of analytical methods within the food laboratory is then discussed.

'Food analysis' is generally understood to refer to chemical or physical tests, assays or measurements, and could include a wide variety of tests, such as determination of water, fat, fibre, nitrite or nitrate content, and measurement of mycotoxins, pesticide or herbicide residues. Microbiological tests usually involve determination of presence or absence of pathogenic microbes, or estimation of numbers of pathogenic, indicator, 'total' numbers or various species of spoilage organism. Analysis of the microbiological quality of food is sometimes referred to as 'food examination' and has been treated separately from food analysis.

In the UK a distinction is made between chemical analysis and microbiological examination for the purposes of the Food Safety Act 1990[1]

and Regulations made under the Act. This is unusual in that most countries do not make this distinction and for them 'analysis' embraces both chemical and microbiological analyses. Thus, in the EU general analysis legislation is taken to refer to both chemistry and microbiology. It is important that this is appreciated when non-UK analytical documents are considered. One reason for this distinction in the UK has been the recognition that it has been considered more difficult to apply quality control systems to microbiological tests. This is because microbes are heterogeneously distributed in many foods, because it is much more difficult to prepare stable control samples, and also perhaps because microbiology has traditionally been considered to some degree an art – the results depending to some extent on the particular skill of individual microbiologists. This has been particularly true for the detection/isolation of pathogens such as salmonellas, which relied on the ability to spot one or a few suspect salmonella colonies among hundreds of non-salmonella competitors, and contrasted with analytical methods which invariably rely on a more objective measurement, such as weight, volume or absorbance.

Although many official laboratories still use traditional (colony-count type) microbiological methods, more rapid methods, often partly or wholely mechanised, more akin to those used in chemical analyses, are gradually gaining popularity. Development of better traditional-type methods including selective media with better indicator systems (e.g. chromogenic substrates), or immunological or PCR-type tests applied after enrichment, have also made microbiological testing less subjective than in the past. In addition, it is becoming increasingly clear that similar quality control systems can and should be applied to microbiological as to analytical tests, even though the variability of the results can be much higher. Chapter 5 deals with most of the microbiological aspects of quality assurance systems. This chapter will outline the legislative aspects of assurance of food laboratory performance, applied in the first instance to chemical analyses, highlighting differences with respect to microbiological examination.

5.2 Legislation and codes of practice

Methods of analysis have been prescribed by legislation for a number of foodstuffs since the UK acceded to the European Community in 1972. However, the Community now recognises that the quality of results from a laboratory is equally as important as the method used to obtain the results. This is best illustrated by consideration of the Council Directive on the Official Control of Foodstuffs (OCF) which was adopted by the Community in June 1989.[2] This, and the similar Codex Alimentarius Commission requirements, are described below. As a result of this general recognition there is a general move away from the need to prescribe analytical methods in detail towards the prescription of the general quality systems within which the laboratory must operate. This allows greater flexibility to the laboratory without detracting from the quality of results that it will produce.

Although such an approach is relatively easy to appreciate in the chemistry area, it is less so in the microbiology area where the experimental result is frequently dependent on the method of analysis used – i.e. it is empirical. It is all the more surprising, therefore, that until comparatively recently, although the EU laid down microbiological criteria in Directives for various foods (e.g. egg products, live and cooked shellfish, milk and milk products, including cheeses of various types, meat preparations and minced meat) the methods to be used were not precisely defined.

In the last few years policy on this has changed, however, and the Member States of the EU have taken the lead in the development of new standard methods and the revision of old ones via CEN (European Committee for Standardisation), ISO (International Organization for Standardisation) and IDF (International Dairy Federation). New and revised standard methods include sections on their repeatability and reproducibility, and there are standards completed and in preparation concerned with general quality systems – e.g. Methods for microbiological examination of food and animal feeding stuffs. General laboratory practices[3] and Guidelines on quality assurance and performance testing of culture media.[4] An EU Regulation on microbiological criteria for foodstuffs is in draft,[5] which will specify ISO standard methods to be used for checking specifications previously laid down. It will also permit alternative methods to be used, that have been validated according to the EN/ISO protocol for method validation.

5.3 Legislation in the EU

5.3.1 Official Control of Foodstuffs Directive (OCF) 1989

The Council Directive on the Official Control of Foodstuffs (OCF) which was adopted by the Community in 1989[2] looked forward to the establishment of laboratory quality standards, by stating that 'In order to ensure that the application of this Directive is uniform throughout the Member States, the Commission shall, within one year of its adoption, make a report to the European Parliament and to the Council on the possibility of establishing Community quality standards for all laboratories involved in inspection and sampling under this Directive' (Article 13).

5.3.2 Additional Measures concerning the Official Control of Foodstuffs (AMFC) Directive 1993

Following that the Commission, in September 1990, produced a Report which recommended establishing Community quality standards for all laboratories involved in inspection and sampling under the OCF Directive. Proposals on this were adopted by the Community in the 1993 Directive on Additional Measures Concerning the Official Control of Foodstuffs (AMFC).[6]

In Article 3 of the AMFC Directive it states:

1. Member States shall take all measures necessary to ensure that the laboratories referred to in Article 7 of Directive 89/397/EEC[2] comply with the general criteria for the operation of testing laboratories laid down in European standard EN 45001[7] supplemented by Standard Operating Procedures and the random audit of their compliance by quality assurance personnel, in accordance with the OECD (Organisation of Economic Co-operation and Development) principles Nos. 2 and 7 of good laboratory practice as set out in Section II of Annex 2 of the Decision of the Council of the OECD of 12 Mar 1981 concerning the mutual acceptance of data in the assessment of chemicals.[8]
2. In assessing the laboratories referred to in Article 7 of Directive 89/397/EEC Member States shall:
   (a) apply the criteria laid down in European standard EN 45002;[9] and
   (b) require the use of proficiency testing schemes as far as appropriate.
   Laboratories meeting the assessment criteria shall be presumed to fulfil the criteria referred to in paragraph 1.
   Laboratories which do not meet the assessment criteria shall not be considered as laboratories referred to in Article 7 of the said Directive.
3. Member States shall designate bodies responsible for the assessment of laboratories as referred to in Article 7 of Directive 89/397/EEC. These bodies shall comply with the general criteria for laboratory accreditation bodies laid down in European Standard EN 45003.[10]
4. The accreditation and assessment of testing laboratories referred to in this article may relate to individual tests or groups of tests. Any appropriate deviation in the way in which the standards referred to in paragraphs 1, 2 and 3 are applied shall be adopted in accordance with the procedure laid down in Article 8.'

and in Article 4, it states:

> Member States shall ensure that the validation of methods of analysis used within the context of official control of foodstuffs by the laboratories referred to in Article 7 of Directive 89/397/EEC comply whenever possible with the provisions of paragraphs 1 and 2 of the Annex to Council Directive 85/591/EEC of 23 December 1985 concerning the introduction of Community methods of sampling and analysis for the monitoring of foodstuffs intended for human consumption.[11]

As a result of the adoption of the above Directives, legislation is now in place to ensure that there is confidence not only in national laboratories but also those of the other Member States. As one of the objectives of the EU is to promote the

concept of mutual recognition, this is being achieved in the laboratory area by the adoption of the AMFC Directive.

In addition it is important that there is dialogue and co-operation by the laboratory with its customers. This is also required by virtue of the EN 45001 Standard at paragraph 6, and will be emphasised even more in future revised versions of EN 45001 and ISO/IEC Guide 25.[12]

This Directive is currently undergoing revision, but it is not expected that the laboratory requirements will be any less stringent than is the current legislation.

5.4 The Codex Alimentarius Commission

The decisions of the Codex Alimentarius Commission (CAC) are becoming increasingly important because of the acceptance of Codex Standards in World Trade Organisation (WTO) agreements. They can be regarded as being semi-legal in status. Thus, on a world-wide level, the establishment of the WTO and the formal acceptance of the Agreements on the Application of Sanitary and Phytosanitary Measures (SPS Agreement) and Technical Barriers to Trade (TBT Agreement) have dramatically increased the status of Codex as a body. As a result, Codex Standards are now seen as *de facto* international standards and are increasingly being adopted by reference into the food law of both developed and developing countries.

Because of the status of the CAC described above, the work that it has carried out in the area of laboratory quality assurance must be carefully considered. One of the CAC Committees, the Codex Committee on Methods of Analysis and Sampling (CCMAS) has developed criteria for assessing the competence of testing laboratories involved in the official import and export control of foods. These were recommended by the Committee at its Twenty-first Session in March 1997[13] and adopted by the Codex Alimentarius Commission at its Twenty-second Session in June 1997;[14] they are intended to assist countries in their fair trade in foodstuffs and to protect consumers. They mirror the EU recommendations for laboratory quality standards and methods of analysis.

The criteria for laboratories involved in the import and export control of foods, now adopted by the Codex Alimentarius Commission are:

- to comply with the general criteria for testing laboratories laid down in ISO/IEC (The International Electrotechnical Commission) Guide 25: 1990 'General requirements for the competence of calibration and testing laboratories';[12] (i.e. effectively accreditation)
- to participate in appropriate proficiency testing schemes for food analysis which conform to the requirements laid down in 'The International Harmonised Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories',[15] (already adopted for Codex purposes by the CAC at its 21st Session in July 1995)
- to use, whenever available, methods of analysis which have been validated according to the principles laid down by the CAC, and

- to use internal quality control procedures, such as those described in the 'Harmonised Guidelines for Internal Quality Control in Analytical Chemistry Laboratories'.[16]

In addition, the bodies assessing the laboratories should comply with the general criteria for laboratory accreditation, such as those laid down in the ISO/IEC Guide 58:1993: 'Calibration and testing laboratory accreditation systems – General requirements for operation and recognition'.[17]

Thus, as for the European Union, the requirements are based on accreditation, proficiency testing, the use of validated methods of analysis and, in addition, the formal requirement to use internal quality control procedures which comply with the Harmonised Guidelines. Although the EU and Codex Alimentarius Commission refer to different sets of accreditation standards, the ISO/IEC Guide 25: 1990 and EN 45000 Series of Standards are similar in intent. It is only through these measures that international trade will be facilitated and the requirements to allow mutual recognition to be fulfilled will be achieved. They both aim to facilitate international trade by enabling mutual recognition of efficient analytical laboratories. However, all of these Standards have effectively been replaced by the ISO/IEC Standard 17025.

5.5 The UK Food Standards Agency

5.5.1 Surveys

The Food Standards Agency undertakes food survey exercises. It has developed information for potential contractors on the analytical quality assurance requirements for food chemical surveillance exercises. These requirements are outlined below; they emphasise the need for a laboratory to produce and report data of appropriate quality. The requirements are divided into three parts dealing with:

Part A: quality assurance requirements for surveillance projects provided by potential contractors at the time tender documents are completed and when commissioning a survey. Here information is sought on:

- the formal quality system in the laboratory if third party assessed (e.g. if UKAS accredited or GLP compliant)
- the quality system if not accredited
- proficiency testing
- Internal Quality Control
- Method Validation

Part B: information to be defined by the FSA customer once the contract has been awarded – to be agreed with contractor, e.g. the sample storage conditions to be used, the methods to be used and a copy of Standard Operating Procedures (SOPs) where accredited, the internal quality control (IQC) procedures to be used, the measurement limits (i.e. limit of detection (LOD): limit of determination/quantification (LOQ); reporting limits and the measurement uncertainty).

Part C: information to be provided by the contractor on an on-going basis once the contract is awarded – to be agreed with the customer to ensure that the contractor remains in 'analytical control'.

5.5.2 Contractual research

The procedures employed by the UK Food Standards Agency have recently been reviewed by a Working Group chaired by Sir John Arbuthnott and his Report published in 2001.[18] That Report recommends that the quality systems employed by the Agency's research contractors be reviewed with a view to the introduction of formal third-party assessed system by 2006.

5.6 Quality assurance requirements: accreditation

The effect of the AMFC (Additional Measures Concerning the Official Control of Foodstuffs) Directive is that organisations must consider the following aspects within the laboratory: its organisation, how well it actually carries out analyses, and the methods of analysis used in the laboratory. All these aspects are inter-related, but in simple terms may be thought of as:

- becoming accredited to an internationally recognised Standard; such accreditation is aided by the use of internal quality control procedures,
- participating in proficiency schemes, and
- using validated methods.

The AMFC Directive requires that food control laboratories should be accredited to the EN 45000 series of Standards as supplemented by some of the OECD GLP principles. In the UK, Government Departments have nominated the United Kingdom Accreditation Service (UKAS) to carry out the accreditation of official food control laboratories for all the aspects prescribed in the Directive. However, as the accreditation agency will also be required to comply to EN 45003 Standard and to carry out assessments in accordance with the EN 45002 Standard, any other accreditation agencies that are members of the European Co-operation for Accreditation of Laboratories (EA) may also be nominated to carry out the accreditation. Similar procedures will be followed in the other Member States, all having or developing equivalent organisations to UKAS. It is the normal practice for UKAS to accredit laboratories on a method-by-method basis although the accreditation of generic protocols (i.e. instrumental procedures) is becoming increasingly more frequent.

In the UK, official food control laboratories undertaking microbiological examination are accredited on a method-by-method basis for the detection and/or enumeration of pathogenic indicators and organisms routinely determined in food, including aerobic colony count, Enterobacteriaceae or coliforms, *E. coli* (including serotype 0157), *Staphylococcus aureus, Bacillus cereus, Clostridium perfringens, Salmonella* species, *Listeria monocytogenes* and *Campylobacter*

species. Where legislation prescribes the methods to be used, any official laboratory which intends to use such methods must be accredited to use them. If not prescribed by statute, then methods published by ISO, CEN, AOAC INTERNATIONAL or other methods which have been validated may be used.

Microbiological examinations for which there are no approved standard methods may be undertaken where the laboratory has in place a series of accredited specific methods or accredited generic protocols dealing with, for example: sample preparation, colony counting, impedimetric techniques, immunological procedures, gene probe methods, PCR and electron or other microscopy techniques. It will be necessary for laboratories to be able to demonstrate quality control procedures to ensure compliance with the EN 45001 Standard,[7] an example of which would be compliance with the ISO/AOACI/IUPAC (International Union of Pure and Applied Chemistry) Guidelines on Internal Quality Control in Analytical Chemistry Laboratories.[16]

For both of these legislative requirements it is the ISO Standard 17025 which is now applicable. This is similar in intent to the ISO/IEC Guide 25[12] and the equivalent EN Standards, but does lay more emphasis on method validation, measurement uncertainty and traceability than did the previous Standards/Guides. Requirements for accreditation of microbiological laboratories are summarised by the European Co-operation for Accreditation (E.A.).[19]

5.7 Internal quality control (IQC)

Although the legislative requirements apply only to food control laboratories, the effect of their adoption is that other food laboratories are advised to achieve the same standard in order for their results to be recognised as equivalent and accepted for 'due diligence' purposes. In addition, the Codex requirements affect all organisations involved in international trade and thus provide an important 'quality umbrella'.

As shown above, these include the requirements for a laboratory to be third-party assessed to international accreditation standards, to demonstrate that it is in statistical control by using appropriate internal quality control procedures, to participate in proficiency testing schemes which provide an objective means of assessing and documenting the reliability of the data it is producing and to use methods of analysis which are 'fit-for-purpose'. These requirements may be summarised as follows and then described in greater detail later in this chapter:

IQC is one of a number of concerted measures that analytical chemists can take to ensure that the data produced in the laboratory are of known quality and certainty. In practice this is determined by comparing the results achieved in the laboratory at a given time with a standard. IQC therefore comprises the routine practical procedures that enable the analyst to accept a result or group of results or reject the results and repeat the analysis. IQC is undertaken by the inclusion of particular reference materials, 'control materials', into the analytical sequence

and by duplicate analysis. IQC for food microbiology is discussed in Chapter 6. ISO, IUPAC and AOAC INTERNATIONAL have co-operated to produce agreed protocols on the 'Design, Conduct and Interpretation of Collaborative Studies'[20] and on the 'Proficiency Testing of (Chemical) Analytical Laboratories'.[15] The Working Group that produced these protocols has prepared a further protocol on the internal quality control of data produced in analytical laboratories. The document was finalised in 1994 and published in 1995 as the 'Harmonised Guidelines For Internal Quality Control In Analytical Chemistry Laboratories' (IQC protocol).[16] The use of the procedures outlined in the Protocol should aid compliance with the accreditation requirements specified above.

Internal quality control in microbiology laboratories differs somewhat from the procedures in analytical laboratories, mostly because the analyte is less stable. However, 'Standard' or 'reference' materials can be used, as can replicate testing and replicate evaluation of test results and the spiking of samples with appropriate standard strains of organisms (see Chapter 6).

5.7.1 Basic concepts

The IQC protocol sets out guidelines for the implementation of internal quality control (IQC) in analytical laboratories. IQC is one of a number of concerted measures that analytical chemists can take to ensure that the data produced in the laboratory are fit for their intended purpose. In practice, fitness for purpose is determined by a comparison of the accuracy achieved in a laboratory at a given time with a required level of accuracy. Internal quality control therefore comprises the routine practical procedures that enable the analytical chemist to accept a result or group of results as fit for purpose, or reject the results and repeat the analysis. As such, IQC is an important determinant of the quality of analytical data, and is recognised as such by accreditation agencies.

Internal quality control is undertaken by the inclusion of particular reference materials, called 'control materials', into the analytical sequence and by duplicate analysis. The control materials should, wherever possible, be representative of the test materials under consideration in respect of matrix composition, the state of physical preparation and the concentration range of the analyte. As the control materials are treated in exactly the same way as the test materials, they are regarded as surrogates that can be used to characterise the performance of the analytical system, both at a specific time and over longer intervals. Internal quality control is a final check of the correct execution of all of the procedures (including calibration) that are prescribed in the analytical protocol and all of the other quality assurance measures that underlie good analytical practice. IQC is therefore necessarily retrospective. It is also required to be as far as possible independent of the analytical protocol, especially the calibration, that it is designed to test.

Ideally both the control materials and those used to create the calibration should be traceable to appropriate certified reference materials or a recognised

empirical reference method. When this is not possible, control materials should be traceable at least to a material of guaranteed purity or other well-characterised material. However, the two paths of traceability must not become coincident at too late a stage in the analytical process. For instance, if control materials and calibration standards were prepared from a single stock solution of analyte, IQC would not detect any inaccuracy stemming from the incorrect preparation of the stock solution.

In a typical analytical situation several, or perhaps many, similar test materials will be analysed together, and control materials will be included in the group. Often determinations will be duplicated by the analysis of separate test portions of the same material. Such a group of materials is referred to as an analytical 'run'. (The words 'set', 'series' and 'batch' have also been used as synonyms for 'run'.) Runs are regarded as being analysed under effectively constant conditions. The batches of reagents, the instrument settings, the analyst, and the laboratory environment will, under ideal conditions, remain unchanged during analysis of a run. Systematic errors should therefore remain constant during a run, as should the values of the parameters that describe random errors. As the monitoring of these errors is of concern, the run is the basic operational unit of IQC.

A run is therefore regarded as being carried out under repeatability conditions, i.e., the random measurement errors are of a magnitude that would be encountered in a 'short' period of time. In practice the analysis of a run may occupy sufficient time for small systematic changes to occur. For example, reagents may degrade, instruments may drift, minor adjustments to instrumental settings may be called for, or the laboratory temperature may rise. However, these systematic effects are, for the purposes of IQC, subsumed into the repeatability variations. Sorting the materials making up a run into a randomised order converts the effects of drift into random errors.

Spiked samples of the food being examined microbiologically can be used. There are a number of difficulties/uncertainties with this approach, however. Firstly, unless the food has been sterilized, it is not possible to be quite sure that it does not already contain the target organism. Secondly, if the food is sterile (e.g. by autoclaving or (preferably) irradiating) then there will be no competitive flora. Thirdly, the wild strains sought may have different properties from the control strain used to spike the food and may also differ in their physiological state. The first difficulty is relatively easily overcome by using a relatively rare strain that can easily be recognised when isolated. This is also a useful precaution in case of accidental cross-contamination from the 'positive control' to the test culture(s). For instance, cultural methods for isolating *Salmonella* species are extremely sensitive, such that even one organism per 25 g sample can be detected with relative ease. The consequences of reporting a sample of processed food positive for *Salmonella* can be extremely serious involving recall of large quantities of product and potential losses of millions of pounds.

5.7.2 Recommendations

Specific recommendations are given in the guidelines which represent integrated approaches to IQC that are suitable for many types of analysis and applications areas. Managers of laboratory quality systems will have to adapt the recommendations to the demands of their own particular requirements. Such adoption could be implemented, for example, by adjusting the number of duplicates and control material inserted into a run, or by the inclusion of any additional measures favoured in the particular application area. The procedure finally chosen and its accompanying decision rules must be codified in an IQC protocol that is separate from the analytical system protocol.

The practical approach to quality control is determined by the frequency with which the measurement is carried out and the size and nature of each run. The use of control charts and decision rules are covered in Appendix 1 to the guidelines.

By following the guidelines laboratories would introduce internal quality control measures which are an essential aspect of ensuring that data released from a laboratory are fit-for-purpose. If properly executed, quality control methods can monitor the various aspects of data quality on a run-by-run basis. In runs where performance falls outside acceptable limits, the data produced can be rejected and, after remedial action on the analytical system, the analysis can be repeated.

The guidelines stress, however, that internal quality control is not foolproof even when properly executed. Obviously it is subject to 'errors of both kinds', i.e., runs that are in control will occasionally be rejected and runs that are out of control occasionally accepted. Of more importance, IQC cannot usually identify sporadic gross errors or short-term disturbances in the analytical system that affect the results for individual test materials. Moreover, inferences based on IQC results are applicable only to test materials that fall within the scope of the analytical method validation. Despite these limitations, which professional experience and diligence can alleviate to a degree, internal quality control is the principal recourse available for ensuring that only data of appropriate quality are released from a laboratory. When properly executed it is very successful.

The guidelines also stress that the perfunctory execution of any quality system will not guarantee the production of data of adequate quality. The correct procedures for feedback, remedial action and staff motivation must also be documented and acted upon. In other words, there must be a genuine commitment to quality within a laboratory for an internal quality control programme to succeed, i.e., the IQC must be part of a complete quality management system.

5.7.3 Quality control of media

Almost all microbiological tests require the use of media, most of which are not chemically defined, but contain mixtures of nutrients, and frequently selective agents, designed to inhibit unwanted microbes, as well as indicator systems

designed to identify colonies of the microbes sought. The proper performance of these media is therefore essential if a laboratory is to obtain reliable results of testing. Although recipes with detailed lists of ingredients and instructions for preparation are provided in standard test protocols, few laboratories prepare their media from basic ingredients. Most buy them in dehydrated form, which need only to be mixed with the correct quantity of water and sterilised. Heat labile ingredients are added after sterilisation of the basal media and are also available commercially ready to use. Medium manufacturers test the functioning of the ingredients of their media (e.g. gelling properties of agar, composition of peptones, inhibitory effect of bile salts, brilliant green) and also the functioning of the complete medium. Nevertheless, laboratories should check the functioning of each new lot of medium they buy. This is normally done by use of test inocula of target and (in the case of selective media) unwanted (competitive) microflora.[21] The choice of test organisms may include recent isolates which reflect ones most likely to be encountered in future, as well as standard strains, and possibly strains known to be particularly sensitive to suboptimal media. Methods of testing can be quantitative (comparison of colony counts on control versus test media) or semi-quantitative (standardised streaking or 'ecometry'). The appearance and size, as well as the number of colonies should be checked. Methods have also been devised for liquid media. Standard methods are in preparation on this topic.[4, 22] Less detailed tests (e.g. a qualitative streak-plate) should also be set up for each batch of medium sterilised.

5.8 Proficiency testing

Participation in proficiency testing schemes provides laboratories with an objective means of assessing and documenting the reliability of the data they are producing. Although there are several types of proficiency testing schemes they all share a common feature: test results obtained by one laboratory are compared with those obtained by one or more testing laboratories. The proficiency testing schemes must provide a transparent interpretation and assessment of results. Laboratories wishing to demonstrate their proficiency should seek and participate in proficiency testing schemes relevant to their area of work. A proficiency testing scheme is defined as a system for objectively checking laboratory results by an external agency. It includes comparison of a laboratory's results at intervals with those of other laboratories, the main object being the establishment of trueness.

In addition, although various protocols for proficiency testing schemes have been produced the need now is for a harmonised protocol that will be universally accepted; the progress towards the preparation and adoption of an internationally recognised protocol is described below. Various terms have been used to describe schemes conforming to the draft protocol (e.g. external quality assessment, performance schemes etc.), but the preferred term is 'proficiency testing'.

Proficiency testing schemes are based on the regular circulation of homogeneous samples by a co-ordinator, analysis of samples (normally by the laboratory's method of choice) and an assessment of the results. However, although many organisations carry out such schemes, there has been no international agreement on how this should be done – in contrast to the collaborative trial situation. In order to rectify this, the same international group which drew up collaborative trial protocols was invited to prepare one for proficiency schemes (the first meeting to do so was held in April 1989). Other organisations, such as CEN, are also expected to address the problem.

5.8.1 Microbiological proficiency testing schemes

Currently there are no internationally or nationally recognised standards for proficiency testing schemes for the microbiological examination of food. The available proficiency testing schemes for food examination will therefore be recognised by the FSA on a case by case basis. Schemes that satisfy the requirements will be recognised and food examination laboratories wishing to be recognised as official control laboratories will be required to participate in the relevant parts of one or more of the recognised schemes. The FSA requires that schemes recognised for the purposes of the AMCF must comply with the general principles of the International Harmonised Protocol for Proficiency Testing of (Chemical) Analytical Laboratories[15] in as far as they are appropriate.

Proficiency test samples should mirror routine situations likely to be encountered when examining foods in the UK under the AMCF. There should be at least 12 distributions per year. Each distribution may contain a number of test materials. Each test material may contain a single organism, a mixture of organisms or may be devoid of organisms of significance. Detection and/or determination of specific pathogenic organisms and indicators are required at least once each year.

Where quantitative determinations are assessed, schemes should treat the results statistically to determine whether performance is satisfactory, for example by converting counts to $\log_{10}$ values and then applying the procedures which have been developed in the International Harmonised Protocol for Proficiency Testing of (Chemical) Analytical Laboratories. Recognised proficiency testing schemes for the microbiological examination of food shall also include a procedure for the recognition of unsatisfactory qualitative results. Currently there are no nationally or internationally recognised protocols for assessing satisfactory performance in qualitative (presence/absence) food examinations. Nevertheless, it is proposed that, in assessing performances, schemes should take due account of false positive and false negative results.

Proficiency test results which fall outside acceptable confidence intervals prescribed for the schemes are unsatisfactory. In such cases it will be necessary for laboratories to demonstrate to UKAS that appropriate remedial action has been taken. The performance of official laboratories in proficiency testing schemes recognised as suitable for official control laboratories will be monitored

by UKAS on behalf of the FSA. Official food control laboratories must therefore consent to UKAS reporting to the FSA on their performance.

5.8.2 Why proficiency testing is important

Participation in proficiency testing schemes provides laboratories with a means of objectively assessing, and demonstrating, the reliability of the data they produce. Although there are several types of scheme, they all share a common feature of comparing test results obtained by one testing laboratory with those obtained by other testing laboratories. Schemes may be 'open' to any laboratory or participation may be invited. Schemes may set out to assess the competence of laboratories undertaking a very specific analysis (e.g. lead in blood) or more general analysis (e.g. food analysis). Although accreditation and proficiency testing are separate exercises, it is anticipated that accreditation assessments will increasingly use proficiency testing data.

5.8.3 Accreditation agencies attitude to proficiency testing

It is now recommended by ISO/IEC Standard 17025,[12] the prime standard to which accreditation agencies now operate, that such agencies require laboratories seeking accreditation to participate in an appropriate proficiency testing scheme before accreditation is gained. There is now an internationally recognised protocol to which proficiency testing schemes should comply; this is the IUPAC/AOAC/ISO Harmonised Protocol. Because of the importance of proficiency testing the Protocol is outlined in the Annex to this Chapter. The elements of the Protocol apply equally to microbiological as well as to chemical measurements.

5.8.4 Blind PT schemes

It should be recognised by laboratories that the use of blind proficiency testing, i.e. where the laboratory receives a sample for analysis from a customer who knows the characteristics of the sample but does not inform the laboratory of that, is becoming more frequent. This is because some customers wish to assess for themselves the effectiveness of their contractors.

5.9 Quality assurance requirements: analytical methods

Methods should be validated as being fit for purpose before use by a laboratory. Laboratories should ensure that, as a minimum, the methods they used are fully documented, laboratory staff trained in their use and control mechanisms established to ensure that the procedures are under statistical control. The development of methods of analysis for incorporation into International Standards or into foodstuff legislation was, until comparatively recently, not

systematic. However, the EU and Codex have requirements regarding methods of analysis and these are outlined below. They are followed by other International Standardising Organisations (e.g. AOAC INTERNATIONAL (AOACI) and the European Committee for Standardization (CEN)).

5.9.1 Codex Alimentarius requirements

This was the first International Organisation working at the government level in the food sector which laid down principles for the establishment of its methods. That it was necessary for such guidelines and principles to be laid down reflects the confused and unsatisfactory situation in the development of legislative methods of analysis that existed until the early 1980s in the food sector. The 'Principles For The Establishment Of Codex Methods Of Analysis'[23] are given below; other organisations which subsequently laid down procedures for the development of methods of analysis in their particular sector followed these principles to a significant degree. They require that preference should be given to methods of analysis the reliability of which have been established in respect of the following criteria, selected as appropriate:

- specificity
- accuracy
- precision; repeatability intra-laboratory (within laboratory), reproducibility inter-laboratory (within laboratory and between laboratories)
- limit of detection
- sensitivity
- practicability and applicability under normal laboratory conditions
- other criteria which may be selected as required.

5.9.2 EU requirements

The EU is attempting to harmonise sampling and analysis procedures in an attempt to meet the current demands of the national and international enforcement agencies and the likely increased problems that the open market will bring. To aid this the Union issued a Directive on Sampling and Methods of Analysis.[11] The Directive contains a technical annex, in which the need to carry out a collaborative trial before it can be adopted by the Community is emphasised.

The criteria to which Community methods of analysis for foodstuffs should now conform are as stringent as those recommended by any International Organisation following adoption of the Directive. The requirements follow those described for Codex above, and are given in the Annex to the Directive.

However, the current draft of the revised Official Food Control Directive states that methods should comply with the following:

> Methods of analysis which are to be considered for adoption under the provisions of this Regulation shall be examined with respect to the

following criteria:

- Accuracy
- Applicability (matrix and concentration range)
- Limit of detection
- Limit of determination
- Precision; repeatability intra-laboratory (within laboratory), reproducibility inter-laboratory (within and between laboratories) but generated from collaborative trial data rather than measurement uncertainty considerations
- Recovery
- Selectivity
- Sensitivity
- Linearity
- Other criteria that may be selected as required.

The precision values referred to in 1(5) shall be obtained from a collaborative trial which has been conducted in accordance with an internationally recognised protocol on collaborative trials (e.g. ISO 5725:1994[24] or the IUPAC International Harmonised Protocol). The repeatability and reproducibility values shall be expressed in an internationally recognised form (e.g. the 95% confidence intervals as defined by ISO 5725:1994 or IUPAC). The results from the collaborative trial shall be published or freely available. Methods of analysis which are applicable uniformly to various groups of commodities should be given preference over methods which apply only to individual commodities.

In situations where methods of analysis can only be validated within a single laboratory then they should be validated in accordance with IUPAC Harmonised Guidelines.

Methods of analysis adopted under this Regulation should be edited in the standard layout for methods of analysis recommended by the International Organisation for Standardisation.

The above provisions are equally applicable to microbiological examination as chemical analyses, for which they were originally developed

5.9.3 Other organisations – CEN and AOACI

There are other International Standardising Organisations, most notably the European Committee for Standardization (CEN) and AOACI, which follow similar requirements. Although CEN methods are not prescribed by legislation, the European Commission does place considerable importance on the work that CEN carries out in the development of specific methods in the food sector; CEN has been given direct mandates by the Commission to publish particular methods, e.g. those for the detection of food irradiation. Because of this some of

the methods in the food sector being developed by CEN are described below. CEN, like the other organisations described above, has adopted a set of guidelines to which its Methods Technical Committees should conform when developing a method of analysis. The guidelines are:

> Details of the interlaboratory test on the precision of the method are to be summarised in an annex to the method. It is to be stated that the values derived from the interlaboratory test may not be applicable to analyte concentration ranges and matrices other than given in annex.
>
> The precision clauses shall be worded as follows:
>
> *Repeatability:* 'The absolute difference between two single test results found on identical test materials by one operator using the same apparatus within the shortest feasible time interval will exceed the repeatability value r in not more than 5% of the cases.
>
> The value(s) is (are): . . .'
>
> *Reproducibility:* 'The absolute difference between two single test results on identical test material reported by two laboratories will exceed the reproducibility value R in not more than 5% of the cases.
>
> The value(s) is (are): . . .'
>
> There shall be minimum requirements regarding the information to be given in an Informative Annex, this being:
>
> Year of interlaboratory test and reference to the test report (if available)
> Number of samples
> Number of laboratories retained after eliminating outliers
> Number of outliers (laboratories)
> Number of accepted results
> Mean value (with the respective unit)
> Repeatability standard deviation (s_r) (with the respective unit)
> Repeatability relative standard deviation (RSD_r) (%)
> Repeatability limit (r) w(with the respective units)
> Reproducibility relative standard deviation (s_R) (with the respective unit)
> Reproducibility relative standard deviation (RSD_R) (%)
> Reproducibility limit (R) (with the respective unit)
> Sample types clearly described
> Notes if further information is to be given

5.9.4 Validation requirements of official bodies

Consideration of the above requirements confirms that in future all methods must be fully validated if at all possible – i.e. have been subjected to a collaborative trial conforming to an International recognised Protocol. In addition

this, as described above, is now a legislative requirement in the food sector of the European Union. The concept of the valid analytical method in the food sector, and its requirements, is described below.

5.10 Criteria for valid methods of analysis

It would be simple to say that any new method should be fully tested for the criteria given above. However, the most 'difficult' of these is obtaining the accuracy and precision performance criteria.

5.10.1 Accuracy

Accuracy is defined as the closeness of the agreement between the result of a measurement and a true value of the measurand.[25] It may be assessed by the use of reference materials. However, in microbiological analysis, there is a particular problem. In many instances the numerical value of a characteristic (or criterion) in a Standard, or whether the organism is present or not, is dependent on the procedures used to ascertain its value. This illustrates the need for the [sampling and] analysis provisions in a Standard to be developed at the same time as the specification in the Standard is negotiated to ensure that the characteristics are related to the methodological procedures prescribed.

5.10.2 Precision

Precision is defined as the closeness of agreement between independent test results obtained under prescribed conditions.[26] In a standard method the precision characteristics are obtained from a properly organised collaborative trial, i.e. a trial conforming to the requirements of an International Standard (the AOAC/ISO/IUPAC Harmonised Protocol or the ISO 5725 Standard). Because of the importance of collaborative trials, and the resource that is now being devoted to the assessment of precision characteristics of analytical methods before their acceptance, they are described in detail below:

5.10.3 Collaborative trials

As seen above, all 'official' methods of analysis are required to include precision data. These may be obtained by subjecting the method to a collaborative trial conforming to an internationally agreed protocol. A collaborative trial is a procedure whereby the precision of a method of analysis may be assessed and quantified. The precision of a method is usually expressed in terms of repeatability and reproducibility values. Accuracy is not the objective.

Recently there has been progress towards a universal acceptance of collaboratively tested methods and collaborative trial results and methods, no matter by whom these trials are organised. This has been aided by the

publication of the IUPAC/ISO/AOAC Harmonisation Protocol on Collaborative Studies.[20] That Protocol was developed under the auspices of the International Union of Pure and Applied Chemists (IUPAC) aided by representatives from the major organisations interested in conducting collaborative studies. In particular, from the food sector, the AOAC International, the International Organisation for Standardisation (ISO), the International Dairy Federation (IDF), the Collaborative International Analytical Council for Pesticides (CIPAC), the Nordic Analytical Committee (NMKL), the Codex Committee on Methods of Analysis and Sampling and the International Office of Cocoa and Chocolate were involved. The Protocol gives a series of 11 recommendations dealing with:

- The Components That Make Up A Collaborative Trial
- Participants
- Sample Type
- Sample Homogeneity
- Sample Plan
- The Method(s) to be Tested
- Pilot Study/Pre-trial
- The Trial Proper

5.10.4 Statistical analysis

It is important to appreciate that the statistical significance of the results is wholly dependent on the quality of the data obtained from the trial. Data which contains obvious gross errors should be removed prior to statistical analysis. It is essential that participants inform the trial co-ordinator of any gross error that they know has occurred during the analysis and also if any deviation from the method as written has taken place. The statistical parameters calculated, and the outlier tests performed are those used in the internationally agreed Protocol for the Design, Conduct and Interpretation of Collaborative Studies.[20]

5.10.5 Alternative validation procedures

In the microbiology sector there will be an interest in alternative validation procedures, most notably for 'Test Kits'. Such procedures are currently being prepared by both AOAC International and CEN.

5.10.6 Single laboratory method validation

There is concern in the food analytical community that although methods should ideally be validated by a collaborative trial, this is not always feasible for economic or practical reasons. As a result, IUPAC guidelines are being developed for in-house method validation to give information to analysts on the acceptable procedure in this area. These guidelines have recently been published[27] and point readers to protocols/guidelines in the area.

5.11 Method validation through proficiency testing

The prime objective of proficiency testing is to assess the 'quality' of the laboratory. However, in some proficiency testing schemes a significant number of laboratories will use the same method of analysis. This is particularly the situation for microbiology proficiency testing schemes. As a result there are initiatives to develop procedures for the validation of methods of analysis using the results from proficiency testing schemes when this situation occurs.

5.12 Measurement uncertainty for the microbiologist

5.12.1 Introduction

It is increasingly being recognised by both laboratories and the customers of laboratories that any reported analytical result is an estimate only and the 'true value' will lie within a range around the reported result. The extent of the range for any analytical result may be derived in a number of different ways, e.g. using the results from method validation studies or, determining the inherent variation through different components within the method, i.e. estimating these variances as standard deviations and developing an overall standard deviation for the method. There is some concern within the food analytical community as to the most appropriate way to estimate this variability.

5.12.2 Microbiology laboratories

Few laboratories until recently indicated their uncertainty of measurement when quoting test results, even when these were expressed as numbers of colony-forming units. The reasons given for this included that the distribution of microbes in the substrates examined (particularly solids, such as foods) was inherently heterogeneous, microbes are often present in clumps that break up to varying degrees during sampling, mixing, diluting and plating, reference materials with exactly known numbers of microbes cannot be made. This attitude has changed recently, partly because the results of standard tests are sometimes used to assess whether a food complies with statutory microbial limits, and partly as a result of widespread introduction of quality assurance and accreditation systems into microbiology laboratories. As with chemical analysis, overall errors can be estimated by investigating individual errors within the method (weighing, pipetting, etc., bias from different individuals counting the colonies), but the method generally favoured is to estimate overall uncertainty by determining repeatability and reproducibility of the method concerned. Uncertainty is minimised by quality assurance systems that minimise errors within the method (e.g. temperature, time of incubation, weighing, measurement of pH, productivity and selectivity of culture media, accuracy of volume measurement). However, these cannot be completely eliminated, and other sources of uncertainty are inherent – numbers of microbes in replicate samples

are generally distributed according to the Poisson distribution, so there is inherently greater uncertainty than found in chemical testing. Uncertainty of measurement can be estimated by replicate testing done within the laboratory as well as from results obtained by participation in proficiency testing schemes. Uncertainty will be affected by factors such as the food substrate being tested, and the method used, so needs to be determined for each food/method combination. In general, selective colony-count methods have greater uncertainty than non-selective 'total' colony count methods, counts from liquids are less uncertain than from solid foods. There are a number of useful publications and draft standards on this topic.[22,28–35]

5.13 Future trends

For the microbiological laboratory, as for all laboratories, it is likely that the most significant developments will be the need to demonstrate the quality of their work. For survey work that is readily achieved through accreditation. However the requirement to demonstrate quality, possibly through a third-party assessment, is likely to be adopted by the major funding agencies in the UK. Such a requirement would have a major impact on the work of the laboratory.

5.14 References

1. 'Food Safety Act, 1990', HMSO, London, 1990, **16**.
2. EUROPEAN UNION, *Council Directive 89/397/EEC on the Official Control of Foodstuffs*, O.J. L186 of 30.6.1989.
3. ISO 1996 BS 5763-0:1996, ISO 7218:1996 Methods for microbiological examination of food and animal feeding stuffs. General laboratory practices
4. ISO (draft) 2000 Microbiology of food and animal feeding stuffs – guidelines on quality assurance and performance testing of culture media. Part 1: General guidelines on quality assurance of culture media in the laboratory. DD ENV ISO 11133-1: 2000.
5. EU MICROBIOLOGICAL CRITERIA PROPOSAL.
6. EUROPEAN UNION *Council Directive 93/99/EEC on the Subject of Additional Measures Concerning the Official Control of Foodstuffs*, O.J. L290 of 24.11.1993.
7. EUROPEAN COMMITTEE FOR STANDARDIZATION, *General Criteria for the Operation of Testing Laboratories – European Standard EN 45001*, Brussels, CEN/CENELEC, 1989.
8. ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT, *Decision of the Council of the OECD of 12 Mar 1981 concerning the mutual acceptance of data in the assessment of chemicals*, Paris, OECD, 1981.
9. EUROPEAN COMMITTEE FOR STANDARDIZATION, *General Criteria for the Assessment of Testing Laboratories – European Standard EN45002*, Brussels, CEN/CENELEC, 1989.
10. EUROPEAN COMMITTEE FOR STANDARDIZATION, *General Criteria for Laboratory Accreditation Bodies – European Standard EN45003*, Brussels, CEN/CENELEC, 1989.

11. EUROPEAN UNION, *Council Directive 85/591/EEC Concerning the Introduction of Community Methods of Sampling and Analysis for the Monitoring of Foodstuffs Intended for Human Consumption*, O.J. L372 of 31.12.1985.
12. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *General Requirements for the Competence of Calibration and Testing Laboratories – ISO/IEC Guide 25,* Geneva, ISO, 1990.
13. CODEX ALIMENTARIUS COMMISSION, *Report of the 21st Session of the Codex Committee on Methods of Analysis and Sampling – ALINORM 97/23A*, Rome, FAO, 1997.
14. CODEX ALIMENTARIUS COMMISSION, *Report of the 22nd Session of the Codex Alimentarius Commission – ALINORM 97/37,* Rome, FAO, 1997.
15. INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY, *The International Harmonised Protocol for the Proficiency Testing of (Chemical) Analytical Laboratories*, ed. Thompson M and Wood R, *Pure Appl. Chem.*, 1993 65 2123-2144 (Also published in *J. AOAC International*, 1993 **76** 926–40).
16. INTERNATIONAL UNION OF PURE AND APPLIED CHEMISTRY, *Guidelines on Internal Quality Control in Analytical Chemistry Laboratories*, ed. Thompson M and Wood R, *Pure Appl. Chem.,* 1995 **67** 649–66.
17. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *Calibration and Testing Laboratory Accreditation Systems – General Requirements for Operation and Recognition – ISO/IEC Guide 58*, Geneva, ISO, 1993.
18. A Review of the Food Standards Agency's Research Portfolio and Research Management Systems, Published by the Food Standards Agency, July 2001.
19. EUROPEAN CO-OPERATION FOR ACCREDITATION, *Accreditation for Microbiological Laboratories EA-04/10 2002* (*http://www.european-accreditation.orgl*).
20. HORWITZ W: 'Protocol for the Design, Conduct and Interpretation of Method Performance Studies', *Pure Appl. Chem*, 1988 **60** 855–64. (Revision published 1995).
21. CORRY, J.E.L., CURTIS, G.D.W. and BAIRD, R.M. 2003 *Handbook of Culture Media for Food Microbiology* (2nd Edition). Progress in industrial microbiology, Volume 37, Amsterdam: Elsevier.
22. ISO (draft) 2002a Microbiology of food and animal feeding stuffs – Guidelines on quality assurance and performance testing of culture media. Part 2: Practical guidelines on performance testing of culture media (ISO/DTS 11133-2: 2002).
23. CODEX ALIMENTARIUS COMMISSION, *Procedural Manual of the Codex Alimentarius Commission – Tenth Edition,* Rome, FAO, 1997.
24. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *Precision of Test Methods – Standard 5725*, Geneva, ISO, 1981 [revised 1986 with further revision in preparation].
25. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *International Vocabulary for Basic and General Terms in Metrology – 2nd Edition*, Geneva, ISO, 1993.
26. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, *Terms and Definitions used in Connection with Reference Materials – ISO Guide 30*, Geneva, ISO, 1992.
27. Harmonised guidelines for single-laboratory validation of methods of analysis, Michael Thompson, Stephen L R Ellison and Roger Wood, *Pure Appl. Chem.*, 2002, **74**(5), 835–55.
28. NMKL (1999) Procedure No 8 Measurement of Uncertainty in Microbiological Examination of Foods. Nordic Committee on Food Analysis, Oslo, Norway (www.nmkl.org).

29. VOYSEY, P.A. and JEWELL, K. 1999 Uncertainty Associated with Microbiological Measurement, Project no 29732, Campden and Chorleywood Food Research Association Review no 15, CCFRA, Chipping Campden, UK.
30. UKAS (United Kingdom Accreditation Service) 2000 The Expression of Uncertainty in Testing Edition 1, UKAS Publication ref: LAB 12, London.
31. ISO (draft) 2001a Milk and milk products – Quality control in the microbiological laboratory Part 1: Analyst performance assessment for colony counts ISO/CD 14461-1| IDF 169-1.
32. ISO (draft) 2001b Milk and milk products – Quality control in the microbiological laboratory Part 2: Determination of the reliability of colony counts of parallel plates and subsequent dilution steps ISO/CD 14461-2|IDF 169-2.
33. ISO (draft) 2002b Microbiology – determination of measurement uncertainty. (doc ISO/TC 34/SC9 N 558).
34. ISO (draft) 2002c Statistics of analytical data Protocol for the establishment of precision characteristics of microbiological quantitative methods by interlaboratory studies, 1st draft, ISO/TC 34/SC 9 N 543.
35. NIEMELÄ, S.I., 2002 Uncertainty of quantitative determinations derived by cultivation of microorganisms. 2nd Edition. Centre for Metrology and Accreditation, Advisory Commission for Metrology, Chemistry Section, Expert Group for Microbiology, Helsinki, Finland, Publication J3/2002.

5.15 Appendix: the ISO/IUPAC/AOAC International Harmonized Protocol for Proficiency Testing of Analytical Laboratories

The International Standardising Organisations, AOACI, ISO and IUPAC have co-operated to produce an agreed 'International Harmonised Protocol for Proficiency Testing of (Chemical) Analytical Laboratories'.[11] That protocol is recognised within the food sector of the European Community and also by the Codex Alimentarius Commission. The protocol makes the following recommendations about the organisation of proficiency testing, all of which are important in the food sector:

5.15.1 Framework

Samples must be distributed regularly to participants who are to return results within a given time. The results will be statistically analysed by the organiser and participants will be notified of their performance. Advice will be available to poor performers and participants will be kept fully informed of the scheme's progress. Participants will be identified by code only, to preserve confidentiality.

The scheme's structure for any one analyte or round in a series should be:

- samples prepared
- samples distributed regularly
- participants analyse samples and report results
- results analysed and performance assessed
- participants notified of their performance

- advice available for poor performers, on request
- co-ordinator reviews performance of scheme
- next round commences.

5.15.2 Organization

The running of the scheme will be the responsibility of a co-ordinating laboratory/organization. Sample preparation will either be contracted out or undertaken in house. The co-ordinating laboratory must be of high reputation in the type of analysis being tested. Overall management of the scheme should be in the hands of a small steering committee (Advisory Panel) having representatives from the co-ordinating laboratory (who should be practising laboratory scientists), contract laboratories (if any), appropriate professional bodies and ordinary participants.

5.15.3 Samples

The samples to be distributed must be generally similar in matrix to the unknown samples that are routinely analysed (in respect of matrix composition and analyte concentration range). It is essential they are of acceptable homogeneity and stability. The bulk material prepared must be effectively homogeneous so that all laboratories will receive samples that do not differ significantly in analyte concentration. The co-ordinating laboratory should also show the bulk sample is sufficiently stable to ensure it will not undergo significant change throughout the duration of the proficiency test. Thus, prior to sample distribution, matrix and analyte stability must be determined by analysis after appropriate storage. Ideally the quality checks on samples referred should be performed by a different laboratory from that which prepared the sample, although it is recognised that this would probably cause considerable difficulty to the co-ordinating laboratory. The number of samples to be distributed per round for each analyte should be no more than five.

5.15.4 Frequency of sample distribution

Sample distribution frequency in any one series should not be more than every 2 weeks and not less than every 4 months. A frequency greater than once every 2 weeks could lead to problems in turn-round of samples and results. If the period between distributions extends much beyond 4 months, there will be unacceptable delays in identifying analytical problems and the impact of the scheme on participants will be small. The frequency also relates to the field of application and amount of internal quality control that is required for that field. Thus, although the frequency range stated above should be adhered to, there may be circumstances where it is acceptable for a longer time scale between sample distribution, e.g. if sample throughput per annum is very low. Advice on this respect would be a function of the Advisory Panel.

5.15.5 Estimating the assigned value (the 'true' result)

There are a number of possible approaches to determining the nominally 'true' result for a sample but only three are normally considered. The result may be established from the amount of analyte added to the samples by the laboratory preparing the sample; alternatively, a 'reference' laboratory (or group of such expert laboratories) may be asked to measure the concentration of the analyte using definitive methods or thirdly, the results obtained by the participating laboratories (or a substantial sub-group of these) may be used as the basis for the nominal 'true' result. The organisers of the scheme should provide the participants with a clear statement giving the basis for the assignment of reference values which should take into account the views of the Advisory Panel.

5.15.6 Choice of analytical method

Participants can use the analytical method of their choice except when otherwise instructed to adopt a specified method. It is recommended that all methods should be properly validated before use. In situations where the analytical result is method-dependent the true value will be assessed using those results obtained using a defined procedure. If participants use a method which is not 'equivalent' to the defining method, then an automatic bias in result will occur when their performance is assessed.

5.15.7 Performance criteria

For each analyte in a round a criterion for the performance score may be set, against which the score obtained by a laboratory can be judged. A 'running score' could be calculated to give an assessment of performance spread over a longer period of time.

5.15.8 Reporting results

Reports issued to participants should include data on the results from all laboratories together with participant's own performance score. The original results should be presented to enable participants to check correct data entry. Reports should be made available before the next sample distribution.

Although all results should be reported, it may not be possible to do this in very extensive schemes (e.g. 800 participants determining 15 analyses in a round). Participants should, therefore, receive at least a clear report with the results of all laboratories in histogram form.

15.5.9 Liaison with participants

Participants should be provided with a detailed information pack on joining the scheme. Communication with participants should be by newsletter or annual

report together with a periodic open meeting; participants should be advised of changes in scheme design. Advice should be available to poor performers. Feedback from laboratories should be encouraged so participants contribute to the scheme's development. Participants should view it as *their* scheme rather than one imposed by a distant bureaucracy.

5.15.10 Collusion and falsification of results

Collusion might take place between laboratories so that independent data are not submitted. Proficiency testing schemes should be designed to ensure that there is as little collusion and falsification as possible. For example, alternative samples could be distributed within a round. Also instructions should make it clear that collusion is contrary to professional scientific conduct and serves only to nullify the benefits of proficiency testing.

5.15.11 Statistical procedure for analysis of results

The first stage in producing a score from a result x (a single measurement of analyte concentration in a test material) is to obtain an estimate of the bias, thus:

$$\text{bias} = x - X$$

where X is the true concentration or amount of analyte.

The efficacy of any proficiency test depends on using a reliable value for X. Several methods are available for establishing a working estimate of $\hat{X}$ (i.e. the assigned value):

In the case of microbiological results, they are log transformed.

Formation of a z-score

Most proficiency testing schemes compare bias with a standard error. An obvious approach is to form the z-score given by:

$$z = (x - \hat{X})/\sigma$$

where σ is a standard deviation. σ could be either an estimate of the actual variation encountered in a particular round ($\tilde{s}$) estimated from the laboratories' results after outlier elimination or a target representing the maximum allowed variation consistent with valid data.

A fixed target value for σ is preferable and can be arrived at in several ways. It could be fixed arbitrarily, with a value based on a perception of how laboratories should perform. It could be an estimate of the precision required for a specific task of data interpretation. σ could be derived from a model of precision, such as the 'Howitzer Curve'.[15] However, while this model provides a general picture of reproducibility, substantial deviation from it may be experienced for particular methods.

5.15.12 Interpretation of *z*-scores

If $\hat{X}$ and σ are good estimates of the population mean and standard deviation then z will be approximately normally distributed with a mean of zero and unit standard deviation. An analytical result is described as 'well behaved' when it complies with this condition.

An absolute value of $z(|z|)$ greater than three suggests poor performance in terms of accuracy. This judgement depends on the assumption of the normal distribution, which, outliers apart, seems to be justified in practice. As z is standardised, it is comparable for all analyses and methods. Thus values of z can be combined to give a composite score for a laboratory in one round of a proficiency test.

The z-scores can therefore be interpreted as follows:

$|z| < 2$ 'Satisfactory': will occur in 95% cases produced by 'well behaved results'

$2 < |z| < 3$ 'Questionable': but will occur in ~5% of cases produced by 'well behaved results'

$|z| > 3$ 'Unsatisfactory': will only occur in ~0.1% of cases produced by 'well behaved results'

5.15.13 Combination of results within a round of the trial

There are several methods of combining the z-scores produced by a laboratory in one round of the proficiency test described in the Protocol. They are:

The sum of scores, SO = Σz
The sum of squared scores, SSZ = Σz^2
The sum of absolute values of the scores, SAZ = $\Sigma|z|$

All should be used with caution however. It is the individual z-scores that are the critical consideration when considering the proficiency of a laboratory.

5.15.14 Calculation of running scores

Similar considerations apply for running scores as apply to combination scores above.