Fera Science Ltd (Fera)

Protocol for Proficiency Testing Schemes

Version 6, April 2017

Part 1 – Common Principles

© Copyright Fera 2017. All rights reserved.

PREFACE

This Protocol is a series of inter-related documents. Part 1, this document, sets out an overview of, and the principles common to, all of the PT schemes provided by Fera Science Ltd (Fera). Subsequent parts give scheme specific details. It follows that neither Part 1, nor any of the other parts, can be used in isolation. Part 1 must always be read in conjunction with a scheme specific supporting part and vice versa.

VERSION HISTORY

This Protocol was completely revised in 2009, superseding all proficiency testing scheme Protocols previously published by Fera in any of its incarnations.

Version 6 of April 2017, this version, supersedes Version 5 of September 2016. The changes are as follows:

Update organisation nomenclature throughout

Section 2.4, removal of reference to legacy accreditation

Section 3.6.2, update to reference

Appendix I, glossary, updates to nomenclatures

CONTENTS

1. INTRODUCTION

Fera was vested on 1 April 2015. Fera provides a wide range of proficiency testing (PT) schemes under the brand name of Fapas®. The management of these PT schemes is the sole task of one team within Fera, known internally as the Proficiency Testing Group (PTG).

FAPAS was created in 1990 and was an acronym for Food Analysis Performance Assessment Scheme. Historically, the other PT schemes run by PTG were known by other acronyms, the Food Examination Performance Assessment Scheme (FEPAS), Laboratory Environmental Analysis Proficiency scheme (LEAP) and Genetically Modified Materials Analysis performance scheme (GeMMA). These historical brand names retain their association with their scheme descriptors as follows: Fapas® Food Chemistry scheme (FAPAS), Fapas® Food Microbiology scheme (FEPAS), Fapas® GM scheme (GeMMA), and Fapas® Water and Environmental scheme (LEAP).

This Protocol, Part 1, should be read in conjunction with the scheme-specific parts. For Fapas® Food Chemistry scheme (FAPAS) in its entirety, see also Part 2 of the Protocol. For Fapas® Food Microbiology scheme (FEPAS), see also Part 3. For Fapas® GM scheme (GeMMA), see also Part 4. For Fapas® Water and Environmental scheme (LEAP), see also Part 5.

For the purpose of this Protocol we use Fapas® to mean Fera PTG.

1.1. What is PT?

ISO/IEC 17043:2010 [1] defines PT as the evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons.

The demand for independent demonstration of competence, from regulatory bodies and customers, means that proficiency testing is relevant to all laboratories testing samples for quality and safety. Hence, it is a requirement of accreditation to ISO/IEC 17025 [2] that the laboratory takes part in a PT scheme, if a suitable scheme exists. In particular, for laboratories entrusted with the official control of food and feeds, Article 12 of EU Regulation (EC) 882/2004 [3] requires such laboratories to be assessed and accredited in accordance with ISO/IEC 17025. This is reinforced internationally under Codex guidelines [4]. PT is an important requirement of the EU Council Directive 98/83/EC [5] on the quality of water intended for human consumption. With the increasing demands for independent proof of competence from regulatory bodies and customers, proficiency testing is relevant to all laboratories testing water for quality and safety in every country. Proficiency testing is therefore a legal requirement for these laboratories. Thus, together with the use of validated methods and internal quality control, proficiency testing is an essential element of laboratory quality assurance.

The analysis of an external quality check sample as part of a laboratory's routine procedures provides objective standards for individual laboratories to perform against and permits them to compare their analytical results with those from other laboratories. In summary, PT is a way of checking the accuracy [6] of results from laboratories.

1.2. Accreditation and PT

Accreditation is a completely separate concept to PT. Accreditation requires the formal, external, assessment of an organisation's documented procedures against a relevant International Standard.

The relevant conformity standard for laboratories in the field of testing is ISO/IEC 17025:2005. Compliance with ISO/IEC 17025 alone cannot guarantee that the procedures give accurate results. Only the external check of a proficiency test can confirm that the results are accurate – hence the requirement within ISO/IEC 17025 for laboratories to take part in PT schemes.

It must be stressed that taking part in a PT scheme does not confer accreditation upon a laboratory. This applies even if the PT provider is, as is Fapas®, accredited for the provision of PT schemes.

1.3. Selection of PT by Users

Guidance documents exist [7, 8] for the selection of PT schemes and the level and frequency of participation. In addition, stakeholders (such as accreditation bodies, customers, regulatory authorities) may have specific requirement for laboratories to take part in some PTs. Fapas® provides a very wide range of analyte/matrix combinations throughout the annual programmes. Some PTs will also vary in relation to the concentrations of analytes within the matrix. In general, the PTs are not method-specific. It is the responsibility of the participant to ensure that the selected PT is suitable for their purpose and to contact Fapas® if there is any doubt about its suitability.

2. ORGANISATION OF SCHEMES

2.1. Administration

All PT schemes provided by Fapas® are administered in keeping with internationally agreed principles, in particular those set out within the IUPAC International Harmonized Protocol for the Proficiency Testing of Analytical Chemistry Laboratories [9]. The original (1994) version of this International Harmonized Protocol was derived from the entire first (1991) Fapas® Protocol while the recent revision (2006) drew heavily upon the experience of Fapas® in delivering PTs in the intervening years.

Fapas® maintains an Advisory Group, which meets at least annually. The Advisory Group members comment upon the relevant programme of PTs planned by Fapas® for the forthcoming year and discuss any scientific issues arising from PTs conducted in the current year. Committee members are available to advise Fapas® staff at any point during the year and group email correspondence is frequently used to facilitate discussions. A list of current Advisory Group members and the terms of reference are available on request from Fapas®.

The day to day running of an individual PT is the responsibility of a designated member of staff, the 'Round Co-ordinator'. Ultimate responsibility for all Fapas® PTs lies with the Head of Group. Expert advice to support all staff in these duties is readily available from within Fera and from a variety of external sources. External advisors are selected on the basis of their personal expertise and not their affiliation; they need not be members of the Advisory Group. When consulting experts, Fapas® will not disclose any participant information, purely scientific information will be exchanged, see below.

2.2. Confidentiality

All information held by Fapas® about participants, including their z-scores, is confidential and will not be disclosed to anyone unless explicitly agreed by the participant for a particular purpose. To preserve this confidentiality participants receive reports giving all the results for that PT but without identifying individual laboratories. The laboratory code numbers used in reports are assigned in order of receipt of results from participants. Participants will be assigned the same code number in different PTs only by chance.

To avoid any conflict of interest / breach of confidentiality, if any of the various analytical testing teams elsewhere within Fera wished to participate in a PT they will be treated in exactly the same manner as any other participant. They will not have access to details of any other participants. Likewise, when Fapas® seeks expert advice from other parts of Fera (or indeed any external source) it will not disclose any information that would breach participant confidentiality.

All PT reports issued by Fapas® are copyright Fera. Anyone wishing to use data from within Fapas® reports for their own publications should first seek permission from Fapas®. It should be noted that this request for respect of copyright cannot preclude publications exploiting Fapas® data being distributed without the prior knowledge or approval of Fapas®.

2.3. Typical Timetable

Fapas® provides on-going PT schemes, where test materials are distributed on a regular basis every year. Fapas® also provides bespoke 'closed' PT schemes, where the test materials are distributed at the time and request of a commissioning client.

For ease of planning and timetabling, Fapas® advertises the on-going schemes in annual blocks, from 1 January to 31 December within calendar year. These annual programmes of proficiency tests are compiled by Fapas® in conjunction with the Advisory Group. They are generally published in August in anticipation of the following January-December. Where short date formats are published, the UK convention of DD/MM/YYYY is employed.

The outline process of conducting a single proficiency test is as follows:

a) Preparation of test materials, including homogeneity testing.

b) Dispatch of test materials on the advertised date from Fapas®, York, UK.

c) Participants analyse test materials and report results by a given date. Generally, the closing date is six to eight weeks from the dispatch date, though for certain analyses where the analyte/matrix combinations potentially are unstable a much shorter time scale may be set.

d) Results subjected to statistical analysis by Fapas®.

e) Distribution of final report to all participants. Generally the report is issued within 15 days of the PT closing date (within 25 days for microbiology, GM, water and environmental PTs) but Fapas® reserves the right to extend this period in cases where the statistical evaluation proves to be atypical.

Participants will be kept informed by email if a delay arises at any of these stages.

2.4. Management System

The quality management system for the whole of Fera is certified to ISO 9001 [10]. In addition, the majority of the work of Fapas® is accredited by the United Kingdom Accreditation Service (UKAS) to ISO 17043. The formal accreditation certificate is available on the Fapas® web site [11] (Adobe PDF format), while the current formal schedule detailing the scope of this accreditation can be obtained from the UKAS web site (Adobe PDF format) [12].

The scheme specific supporting parts to this document include the accreditation status of each PT scheme. Where applicable, Fapas® is a UKAS accredited Proficiency Testing Provider, No. 0009.

2.5. Subcontractors

Fapas® does not have any laboratory facilities of its own. Test material preparation, homogeneity testing and stability testing is carried out by subcontractors. Homogeneity testing may be carried out by a different laboratory to the one that prepares the test material. Fapas® maintains a list of approved subcontracting laboratories and regularly reviews the service received. Where possible, Fapas® will only use subcontracting laboratories that hold accreditation to recognised international standards (ISO/IEC 17025 [2], for example). Subcontracting laboratories may also participate in Fapas® PTs. In this situation, the subcontracting laboratory participation will be treated in exactly the same way as all the other participants, and the same rules of confidentiality will apply.

2.6. Agents

Agents are appointed by Fapas® in some countries. The advantages to participants of using the agent are to register locally to participate in Fapas® PTs and the facility to pay in local currency. Agents will also liaise with Fapas® on the participant's behalf for any queries or problems. Agents may also be able to help samples pass more easily through customs. Details of participants' performance in the PTs are not disclosed to the agents. The list of agents is available from the website, www.fapas.com.

2.7. Bespoke schemes

There is the possibility of producing a bespoke PT for customers for which the scheduled PTs do not properly fulfil their purpose. Customers requiring a bespoke service should bear in mind the statistical and economic advantages of including as many participants as possible in a PT. Bespoke PTs use the same Fera Standard Terms and Conditions for Proficiency Testing Schemes as our scheduled PTs. Customers requiring a bespoke service should discuss the specifications in detail with Fapas® in the first instance. Bespoke services may or may not fit within the scope of accreditation.

3. PARTICIPATION IN SCHEMES

None of the Fapas® schemes stipulate a minimum level and frequency of participation. Participants do not necessarily have to analyse for all the analytes in a test.

3.1. Enrolment and Fees

The programmes for Fapas® PTs are available on the web site, www.fapas.com. Customers place their orders on-line by browsing these programmes and compiling a 'wish list'. If the customer is a previous participant and has access to the secure pages of our web site they can convert their 'wish list' into a formal order on-line. New customers can use the 'wish list' to request a quote. Alternatively, PDF files of the programmes are available from Fapas®, at the address shown on the final page of this document.

PT order confirmations are automatically emailed to customers on completion of the ordering process. The confirmation email contains a link to a printer friendly version of the order, held within the customer's secure pages on our web site. It is the responsibility of the customer to check that Fapas® has processed their requests correctly, i.e. that they are enrolled in the correct PTs.

Details of all fees are available on request. Fapas® reserves the right to withhold test materials and/or PT reports from participants if payment is delayed.

Formal Fera Standard Terms and Conditions for Proficiency Testing Schemes are available from our web site (PDF format) www.fapas.com.

3.2. Dispatch and Receipt of Test Materials

All test materials are distributed with a generic compliments slip. The compliments slip provides details on how to access instructions from our website about reporting of results and method details. Instructions specific to the PT with regard to storage on receipt, type of analysis required, etc. will be included in these instructions. The instructions on storage of samples after receipt are advisory, not absolutely fixed. A small range of temperature around the advised storage temperature can be considered to be acceptable.

It is the responsibility of participants to read these instructions and follow them. Fapas® cannot be held responsible for any problems arising from failure to comply with these directions.

It is the responsibility of the participant to contact Fapas® if they have not received the test material within agreed timescales, as set out in the Fera Standard Terms and Conditions for Proficiency Testing Schemes.

Delays to the dispatch of test materials occasionally arise. If the dispatch of a test material has to be delayed for any reason, then participants will be notified of this fact by email prior to the advertised dispatch date. Fapas® cannot be held responsible if participants overlook this notice of delay.

3.2.1. Test material preparation and homogeneity testing

The determinands in test materials may either be at natural levels, incurred or spiked at a particular requested formulation level. Details of test material preparation are retained by Fapas® but not published in PT Reports, except where pertinent to the statistical analysis of the results.

Test materials in Fapas® PTs will not be distributed until testing demonstrates that the individual subsamples are of sufficient homogeneity. Fapas® uses the statistical procedure developed by Fearn and Thompson [13]. Details of the homogeneity testing data are retained by Fapas® but not published in PT Reports.

Participants may contact Fapas® to request details of test material preparation and homogeneity testing. Such details may be released on request, except where this compromises data which is commercial in confidence or where such knowledge is scientifically invalid in the interpretation of assessments.

3.2.2. Stability of test materials

Fapas® proficiency test materials are sufficiently stable for the duration of the test. This includes the time between their preparation and the start of the test, as well as during transportation of test materials and for the period of time set for participants to analyse them. This stability has been established through a combination of formal stability testing, experience of running PTs over more than 25 years, expert advice and assessment of historic data comparing preparation and homogeneity data with the assigned values.

Fapas® reference materials undergo formal stability testing for long-term storage. Data generated for reference materials is also then used to support PT material stability. Additional stability studies have been published by Fapas® for selected test materials [14].

Some PTs will have a reduced timescale where instability over a long time might affect analytical results.

3.2.3. Stability under transportation conditions

Some test materials will be transported from their chilled state in insulated packaging with ice blocks to ensure their continued stability. Some test materials will be shipped in insulated packaging which might be different to the advised storage conditions. It is the responsibility of participants to follow the storage conditions on receipt of test materials as advised on the instruction letter, regardless of the transportation conditions.

3.3. Analysis of Test Materials

If the PT is to yield maximum benefit as an external check on the routine working of participants' methods then the sample should not be given any special treatment. Hence, participants are free to use whatever method of analysis they wish. On the occasions where the method is known to be empirical (i.e. the result is dependent on technique) participants are still free to use whatever method they wish. In order to obtain a comparable set of results for statistical assessment, however, Fapas® may advise participants that only the results submitted for a given method will be used to derive an assigned value by consensus.

3.4. Submission of Results and Methods

The reporting of results within the requested time scale and in the specified units is part of the performance assessment.

Participants are requested to submit their results and methods via the secure pages on our web site. Each participant will confidentially be provided with a unique UserID and Password required to access these pages. While the submission of a result is a prerequisite for a performance assessment, participants are not obliged to submit their method information. However, where an assigned value derived by consensus is dependent on a particular aspect of the method, some specific questions may be required with the result submission.

Acceptance, or otherwise, of results submitted after the closing date is at the discretion of the Round Co-ordinator. Where extenuating circumstances have prevented timely results submission, participants should contact Fapas® to discuss acceptance of late results.

3.4.1. Collusion and Falsification of Results

Collusion, either between participants or between individual participants and the scheme provider, is contrary to professional scientific conduct. It serves only to nullify the benefits of proficiency testing to customers, accreditation bodies, and analysts alike. Collusion is, therefore, to be strongly discouraged.

As a preventive measure Fapas® reserves the right to distribute more than one test material within a PT so that participants cannot compare results directly. Ultimately, though, it is the responsibility of the participating laboratories to avoid collusion or falsification of results. Laboratories found to be falsifying results may be refused participation in subsequent proficiency tests.

3.5. Report Distribution

Participants are advised in the PT instructions when to expect the publication of the report. Fapas® aims to do this as soon as is practical after the closing date of the PT. Participants should note that our quality procedures involve extensive cross-checking and scrutiny by several Fapas® staff under the guidance of the Round Co-ordinator. There is no fixed way of generating an assigned value. Consequently, this means that the process can take a few weeks depending on the complexity of the data.

All reports are distributed in Adobe PDF format. They are both password secured and digitally signed to ensure that they cannot be altered in any way. The digital signature automatically validates when the PDF file is opened using Adobe Reader v7 or higher on a PC with access to the Internet. Reports are only available for download to the named contact(s) for the PT in question.

3.6. Follow-Up Services

If a participant wishes to obtain advice on any aspect of their performance they should contact Fapas® by email in the first instance. All technical queries will be addressed internally initially by the Round Co-ordinator and/or the senior scientist. Participants must note that Fapas® might offer assistance in the form of a broker service whereby Fapas® will either anonymously or subsequent to all parties agreeing to waive their confidentiality, pass on the participant's enquiry to an expert laboratory or external advisor.

3.6.1. Quality Control samples

Surplus test materials from the batch used for the PT may be available for purchase as quality control (QC) samples. These QC samples may be used for troubleshooting poor performance in the PT, training of new staff, method development or generating QC charts. Fapas® QCs have an associated datasheet which provides the same data of assigned values and performance limits as in the PT (see performance assessment sections below). The stability assessment of QC samples is the same as that for the PT.

Certified Reference Materials (CRMs) for the food analysis sector are not numerous and surplus Fapas® test materials may be the only source of a suitable quality control material.

The exact stock level of any given QC can be checked via our web site.

3.6.2. Reference Materials

Reference Materials (RMs) are also available from Fapas® for some analyte/matrix combinations. RMs have a much higher degree of characterisation than PT or QC materials with a defined chain of traceability. RMs undergo formal stability testing for both short-term and long-term applications. RMs have an associated datasheet which lists the reference values and their expanded uncertainty *U*. The value of *U* is not a performance limit but is the uncertainty relating to the reference value. RMs therefore have a greater degree of trust in their values and can be used for method calibration purposes.

Fapas® RMs are generated according to the principles of ISO 17034:2016 [15] and ISO Guides 31 and 35 [16, 17].

The exact stock level of any given RM can be checked via our web site.

4. PERFORMANCE ASSESSMENT

The statistical model used by Fapas® is set out fully within the International Harmonized Protocol [9]. In summary, as indicated in the Introduction, the purpose of a Fapas® PT is to check the accuracy of results submitted by the participating laboratories. This check is achieved typically by comparing participants' results to some estimate of the 'true' value.

If the results submitted are **quantitative** then this comparison will be in the form of a numerical score. Semi-quantitative (< or >) data are not assessed, *except* where detailed in the relevant scheme specific supporting parts of this Protocol.

The comparison for **qualitative** results will be against the answer anticipated by formulation or by taking account of the consensus of participants' results.

The results submitted to a single PT represent the final product in a complex string of actions carried out by the participants, from sample receipt to results reporting. As such they encompass all aspects of a laboratory's performance. A mistake, however trivial, at any stage will contribute to the final outcome.

It is unwise to view any performance assessment as anything other than a snapshot of the whole laboratory performance at the time of the PT.

4.1. Scoring

4.1.1. Why score?

The advantages of expressing participants' results as a standardised score are that:

- they are simple and transparent,
- they present participants' results in a readily understood form,
- they permit comparison over time,
- when tabulated and charted, they place individual performance in the overall context of the PT.

When the standardised score incorporates a prescribed value that represents limits of acceptable variation for the analysis in question then the score embodies the concept of fitness-for-purpose, i.e. the balance between expending considerable time and effort (= expense) on an analysis to get a highly accurate result vs. carrying out a rough and ready procedure that only provides an indication of the level present and so be of limited use/require further analysis.

4.1.2. Types of scores

A variety of standardised scores is available [21]. This Protocol presents only one such score but this does not preclude Fapas® from adopting alternatives, if so advised by our statistical experts.

4.1.2.1. z-Scores

Fapas® favours the use of z-scores because when the standard deviation is based on a fit-for-purpose criterion, i.e. it is a prescribed 'standard deviation for proficiency', then the significance of the performance assessment is immediately apparent, no matter what the concentration or identity of the analyte, the nature of the test material or the physical principle underlying the analytical measurement. By assessing a participant's performance by way of a z-score, both the trueness and the precision of their result are addressed. Use of an objective, fit-for-purpose standard deviation for proficiency requires the measurement uncertainty of a participant's result to be in keeping with this level.

A z-score combines an estimate of the error of a result with a standard deviation:

$$z = \frac{(x - x_a)}{\sigma_p}$$

where x = the result reported by the participant
x_a = the assigned value
and σ_p = the standard deviation for proficiency

The derivation of the assigned value and the choice of fit-for-purpose standard deviations for proficiency are more complex (see sections 4.1.3 and 4.1.4 below). The report for each PT will give full details on the choice and calculation of both the assigned value and the standard deviation for proficiency assessment.

NB. The international standard for statistics in proficiency testing, ISO 13528 [21] was revised in 2015 (the previous version was 2005). The revised standard adopts slightly different nomenclature for the assigned value and standard deviation for proficiency. Fapas® has retained its existing nomenclature for historical consistency [9] and the definitions of the terms are unchanged.

4.1.2.2. Other types of scores

The revised standard ISO 13528:2015 describes the use of alternative scores to the z-score. These include $D_i\%$ (previously adopted by Fapas® as the Q-score but no longer in common use), z' score, zeta (ζ) score and E_n score. The $D_i\%$ provides only the relative error of a result and might have a use where an appropriate standard deviation for proficiency cannot be set.

The z' score combines σ_p with the uncertainty of the assigned value. The zeta score replaces σ_p with the combined uncertainty of the participant's result and the assigned value. The E_n score replaces σ_p with the combined expanded uncertainty of the participant's result and the assigned value. The zeta and E_n scores depend on the participants submitting accurate measurement uncertainty estimates with their result, a procedure not easily adhered to.

Fapas® does not subscribe to the use of the z' score. The value of σ_p in itself is defined by the fitness-for-purpose (see section 4.1.4). The use of the z' score therefore deviates from that pre-defined fitness-for-purpose and is effectively hiding the high uncertainty of the assigned value. Fapas® prefers to issue 'information only' z-scores where the uncertainty is higher than ideal or to not issue z-scores at all where the uncertainty is unacceptable. Further information on the uncertainty of the assigned value is provided below (section 4.1.3). Further information on the use of other scores is provided in Technical Brief 74 [18].

4.1.2.3. Comparing scores across different PTs

Performance assessment scores from different PTs can only be compared if the same rules of fitness-for-purpose are applied. It follows, therefore, that z-scores from different PT providers can only be compared where the value of σ_p has been derived in the same way. Similarly, z-scores and z' scores cannot be compared with any meaning.

4.1.3. Consensus assigned value

In all Fapas® PTs, the 'assigned value', x_a , is the best estimate available to Fapas® of the 'true' value. The assigned value can be set as a:

- consensus value
- formulation level
- certified reference value

Suitable algorithms for the derivation of a consensus value are readily available [19, 20, 21]. A consensus value is almost invariably taken by Fapas® as the assigned value. The procedure used to derive the consensus will involve, as a minimum:

- removing invalid data, i.e., results reported as approximately 10, 100 or 1000x greater or smaller than the majority of submitted results (considered to be reporting errors).
- considering the symmetry, or otherwise, of the distribution of results.
- where the results form a roughly symmetric distribution (outliers aside), minimising the influence of outliers by the use of a robust statistical procedure to derive the mean [19].
- where there is a degree of asymmetry, scrutinising the results with a procedure that estimates the mode or, in some instances, helps to identify multimodality (by a procedure known as 'bump-hunting' [20]).
- comparing the robust mean, median and mode(s). The median or mode may be used as the consensus if Fapas® considers that sufficient supporting evidence is available to justify such action.

Additional procedures may be adopted for particular PTs when results have to be submitted with supporting information, for example, on recovery correction or method-critical parameters. This will be detailed in each specific PT report.

An estimate of the uncertainty of the consensus is also required. For n results, the uncertainty u of a robust mean is taken as its standard error,

$$u = \hat{\sigma} / \sqrt{n}$$

where $\hat{\sigma}$ is the robust standard deviation of the results. For a mode, the uncertainty is taken as the standard error of the mode (SEM), calculated directly by the bootstrap method [20]. For a median, the uncertainty is taken as the median absolute deviation (sMAD) divided by $\sqrt{n}$. Where the test of u/σ_p is equal to or greater than the critical value of 0.4, the effect of the uncertainty on z-scores will be taken into consideration when issuing z-scores. Where the uncertainty is too high, z-scores may be issued for *information only* and should not be used by participants as fully evaluative of performance.

The statistics for the derivation of the assigned value will be summarised in each PT report. Reports will detail any complications in the derivations, as necessary.

4.1.4. Standard deviation for proficiency

The standard deviation for proficiency (informally, the 'target sd', σ_p) determines the limits of performance in a PT. It is set at a value that reflects fitness-for-purpose for the analysis in question. Fit-for-purpose standard deviations for proficiency can be obtained from:

- predictive models, e.g. Horwitz Equation or its modifications [22]
- collaborative trials / method performance studies
- lower limit of interest (LLI)
- legal definition
- expert opinion.

The Horwitz function, describing the trend of standard deviation of reproducibility found in collaborative trials, represents fitness-for-purpose in the food sector over a wide range of concentrations. It is therefore used by Fapas® in the majority of instances. In some ranges, however, a more appropriate precision is required and, in those instances, statistics from relevant collaborative trials or other sources are used to derive the standard deviation. The Horwitz equation may still be used where collaborative trial data exists, if supported by that trial.

The appropriate form of the modified Horwitz equation [22] used by Fapas® requires the analyte concentration c to be expressed as a mass fraction, e.g., $10^{-6} \equiv 1$ ppm $\equiv 1$ mg kg^{-1}, or $10^{-2} \equiv 1\%$. It specifies the following:

For analyte concentrations less than 1.2×10^{-7} (120 ppb),

$$\sigma_p = 0.22c$$

For analyte concentrations between 1.2×10^{-7} (120 ppb) and 0.138 (13.8%),

$$\sigma_p = 0.02c^{0.8495}$$

For analyte concentrations greater than 0.138 (13.8%),

$$\sigma_p = 0.01c^{0.5}$$

Fapas® uses the assigned value x_a as the concentration in these equations. The raw σ_p values are mass fractions and have to be converted to the required units before use in calculating z-scores. This is easily achieved by dividing the result by the mass fraction appropriate to the units used.

Example, sodium in canned meat meal. The robust mean is 0.27 g/100g, where the mass fraction is 10^{-2}. Hence;

$$\sigma_p = 0.02 \times (0.27 \times 10^{-2})^{0.8495} = 0.00013$$

Convert to g/100g units,

$$\sigma_p = 0.00013 / 10^{-2} = 0.013 \text{ g/100g}$$

When collaborative trial statistics are used to determine σ_p , the value at the relevant concentration is obtained by interpolation, using an appropriate model, usually the assumption of a constant relative standard deviation. The function usually applied to derive σ_p is;

$$\sigma_p = \frac{RSD_R}{100} \times c$$

where RSD_R is the Relative Standard Deviation of Reproducibility from collaborative trial, expressed as %

and c is concentration, i.e., the assigned value.

The lower limit of interest (LLI) is used in combination with expert opinion where other standard performance limits (Horwitz or collaborative trial or regulation) are not quite applicable, especially at low concentrations. The LLI generates a correction factor from the low concentration level with an estimate of the relative threshold that applies at this level. The correction factor can then be used directly or applied to the existing standard performance estimator such as the Horwitz equation.

4.2. Interpreting z-Scores

The guiding principle of scoring in Fapas® is fitness-for-purpose. This means that the standard of accuracy required is based on an uncertainty that is independently determined to be appropriate for the analysis in question. A hypothetical laboratory performing exactly according to this predetermined standard will obtain z-scores like a random selection from a normal distribution. However, most laboratories will use methods with both a bias and a repeatability standard deviation that differs from the fitness-for-purpose uncertainty. Accordingly, the collected z-scores from a Fapas® often deviate from the normal distribution. The deviation may take the form of heavy tails and outliers and, occasionally, asymmetry or multimodality. Because the scoring is based on an independently-prescribed uncertainty, it is logical to interpret z-scores on the basis of the normal distribution.

The properties of a normal distribution are such that, over time, about 95% of observations lie between ±2 standard deviations. Performance in a Fapas® PT, therefore, is considered fit-for-purpose if a z-score lies within the range ±2. It follows that an exactly-conforming participant's z-scores will fall outside this range with a probability of 1 in 20. Occasional scores in the range $2 < |z| < 3$ may therefore be of no importance. Such z-scores require consideration and appropriate action, in the context of the other scores obtained by that laboratory. However, the probability of a conforming participant's z-score falling outside $|z| > 3$ is less than about 1 in 300. Given this rarity, such scores therefore represent results that are probably not fit-for-purpose and should be used to trigger investigation and remedial action.

The consideration of a set or sequence of z-scores over time provides more useful information than a single z-score. Examples of suitable methods of comparison are provided in the International Harmonized Protocol [9].

NB. In the past, terms such as 'satisfactory', 'questionable' and 'unsatisfactory' have been applied to z-scores within certain ranges. This approach categorises the z-score when it is not appropriate to do so and is likely to be misleading. The limits $z = \pm 2$, $z = \pm 3$, must not be regarded as strict boundaries but should be treated as action limits. z-Scores are statistics and must be interpreted as such [23].

A note on homogeneity and z-scores: The requirement of distribution units to be 'sufficiently homogeneous' means that any variation detected between the units by the homogeneity test should be of negligible magnitude in relation to fitness-for-purpose and thus too small to affect z-scores. Fapas® therefore takes no account of between-unit uncertainty in its scoring.

4.3. Appeals

Fapas® undertakes to correct any mistakes attributable to errors on its part promptly and sympathetically. If a participant has any concerns about any aspect of the PT they should contact Fapas® by email in the first instance. An investigation will be conducted in accordance with our management system and the participant advised of the outcome.

5. REFERENCES

1 ISO/IEC 17043:2010, Conformity assessment – General requirements for proficiency testing.

2 ISO/IEC 17025:2005, General requirements for the competence of testing and calibration laboratories.

3 Regulation (EC) 882/2004 of the European Parliament and of the Council of 29 April 2004 on official controls performed to ensure the verification of compliance with feed and food law, animal health and animal welfare rules, Official Journal **L165**, 30/04/2004, 0001-0141.

4 CAC/GL 27-1997, Guidelines for the assessment of the competence of testing laboratories involved in the import and export control of food, Codex 2006.

5 Council Directive 98/83/EC of 3 November 1998 on the quality of water intended for human consumption, *Official Journal of the European Communities* L 330/32, 05/12/1998.

6 AMC Technical Brief 13, 2003, Terminology – the key to understanding analytical science. Part 1: Accuracy, precision and uncertainty.

7 EA-4/18: 2010, Guidance on the level and frequency of proficiency testing participation.

8 Selection, Use and Interpretation of Proficiency Testing (PT) Schemes, Eurachem, 2011, second Edition

9 Thompson, M., Ellison, S.L.R. and Wood, R., 2006, The International Harmonised Protocol for the Proficiency Testing of Analytical Chemistry Laboratories, *Pure Appl. Chem.*, **78** (1), 145–196.

10 http://www.fera.co.uk, accessed 17/03/2017

11 http://www.fapas.com, accessed 17/03/2017

12 http://www.ukas.com, accessed 17/03/2017

13 Fearn, T. and Thompson, M., 2001, A new test for sufficient homogeneity, *Analyst*, **126**, 1414-1417.

14 Sykes, M. and Thompson, M., 2015, Assessing the stability of a proficiency test material by participant-blind re-use after a period of storage, *Anal. Methods*, 2015, **7**, 9753

15 ISO 17034:2016, General requirements for the competence of reference material producers

16 ISO Guide 31:2000, Guide to the contents of certificates and labels of reference materials

17 ISO Guide 35:2006, Reference materials – General and statistical principles for certification

18 AMC Technical Brief 74, 2016, z-Scores and other scores in chemical proficiency testing-their meanings, and some common misconceptions, *Anal. Methods*, 2016, **8**, 5553

19 Analytical Methods Committee, 1989, Robust statistics - how not to reject outliers: Part 1 Basic concepts, *Analyst*, **114**, 1693-1697.

20 Lowthian, P.J. and Thompson, M., 2002, Bump-hunting for the proficiency tester-searching for multimodality, *Analyst*, **127**, 1359-1364.

21 ISO 13528:2015, Statistical Methods for use in Proficiency Testing by Interlaboratory Comparison.

22 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.

23 AMC Technical Brief 11, 2002, Understanding and acting on scores obtained in proficiency testing schemes.

APPENDIX 1 GLOSSARY OF TERMS

This glossary includes terms not specifically mentioned in this Protocol but which may be used in the supporting parts of the protocol or PT report. Participants may find this glossary useful in relation to proficiency testing in general.

Accuracy

The closeness of agreement between a test result and the accepted reference value.

NOTE. The term "accuracy", when applied to a set of test results, describes a combination of random components and a common systematic error or bias component.

Assigned value

The value to be used as the "true" value by Fapas® in the statistical treatment of results. It is the best available estimate of the true value of the analyte in the matrix.

Bias

The difference between the expectation of the test results and an accepted reference value.

NOTE. Bias is due to systematic error, not random error. There may be one or more systematic error components contributing to the bias. A larger systematic difference from the accepted reference value is reflected by a larger bias value.

Bias of the measurement method

The difference between the expectation of test results obtained from all laboratories using that method and an accepted reference value.

NOTE. An example of this is where a method purporting to measure the sulfur content of a compound consistently fails to extract all the sulfur, giving a negative bias to the measurement method. The bias of the measurement method is measured by the displacement of the average of results from a large number of different laboratories all using the same method. The bias of a measurement method may be different at different analyte concentrations.

Certified Reference Material (CRM)

A reference material, one or more of whose property values are certified by a technically valid procedure, accompanied by or traceable to a certificate or other documentation which is issued by a certifying body.

Consensus value

The assigned value, as generated from valid participants' results. Some participants' results may be excluded from the consensus calculation where they fail to meet specific criteria. The consensus value may be the robust mean, median or mode.

Distribution unit

One sample of the test material which is sent to a participant.

Error

The difference between a reported result and the assigned value.

Fapas®

Brand name owned by Fera providing the Fapas Food Chemistry (FAPAS), Fapas Food Microbiology (FEPAS), Fapas GM (GeMMA), Fapas Water and Environmental (LEAP) and specialised proficiency testing schemes.

Fera

Fera Science Ltd, parent organisation of Fapas®

Fitness for Purpose

The precision and accuracy of analytical data must be sufficient to enable the end-user of the data to make sound decisions as to whether the results/samples analysed are fit for the intended purpose.

Interlaboratory test comparisons

Organisation, performance and evaluation of tests on the same or similar items or materials by two or more different laboratories in accordance with pre-determined conditions.

Internal Quality Control (IQC)

The set of procedures undertaken by the laboratory staff for continuous monitoring of operations and results in order to decide whether the results are reliable enough to be released; IQC primarily monitors the batch-wise trueness of results on quality control materials, and precision on replicate analysis of test materials.

Laboratory bias

The difference between the expectation of the test results from a particular laboratory and an accepted reference value.

Laboratory component of bias

The difference between the laboratory bias and the bias of the measurement method.

NOTES. (1) The laboratory component of bias is specific to a given laboratory and the conditions of measurement within the laboratory, and it may be different at different analyte concentrations.

(2) The laboratory component of bias is relative to the overall average result, not the true or reference value.

LLI

Lower Limit of Interest

Precision

The closeness of agreement between independent test results obtained under prescribed conditions.

NOTES. (1) Precision depends only on the distribution of random errors and does not relate to the accepted reference value.

(2) The measure of precision is usually expressed in terms of imprecision and computed as a standard deviation of the test results. Higher imprecision is reflected by a larger standard deviation.

(3) “Independent test results” are defined as results obtained in a manner not influenced by any previous result for the same or similar material.

Proficiency Testing Scheme (Performance Assessment Scheme)

The system for objectively checking laboratory results by means of an external agency (e.g. Fapas®). It includes comparison of a laboratory's results at intervals with those of other laboratories, the main object being the establishment of trueness. Proficiency testing is designed to assess the accuracy of a laboratory's results. Proficiency testing is sometimes referred to as "external quality assessment" (EQA).

QC materials

Surplus test materials from the batch used for a PT. Useful for internal quality control (QC) in a laboratory but these are not CRMs.

Quality Assurance System/Programme (QAS)

The sum total of a laboratory's activities aimed at achieving the required standard of analysis. While IQC and proficiency testing are very important components of a quality assurance programme it must also include staff training, administrative procedures, management structure, auditing, etc.

Accreditation bodies judge laboratories on the basis of their quality assurance programme plus peer review of technical competence for a specific technical capability.

Reference Material (RM)

A material or substance one or more properties of which are sufficiently homogeneous and well-established to be used for the calibration of an apparatus, the assessment of a measurement method, or for assigning values to other materials.

Relative Standard Deviation (RSD) / (Coefficient of Variance)

The standard deviation expressed as a percentage of the mean:

$$RSD = \frac{\sigma}{\bar{x}} \times 100$$

where σ is the standard deviation and $\bar{x}$ is the arithmetic mean

Robust mean

The mean of results calculated by a robust statistical method, for example Huber's H15 algorithm as used by Fapas®.

Standard deviation for proficiency (target sd)

A numerical value for the standard deviation of a measurement result, which has been designated as a goal for measurement quality.

Test material

The matrix/analyte combination to be tested that is distributed to participants in the proficiency test.

Test method

A defined technical procedure to determine one or more specified characteristics of a material or product.

Testing laboratory

A laboratory that measures, examines, tests, calibrates or otherwise determines the characteristics or performance of materials or products.

True value

The actual concentration of the analyte in the matrix. Very often, the true value is unknown.

Trueness

The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value.

NOTE. The measure of trueness is usually expressed in terms of bias.

Fapas®
Fera Science Ltd (Fera)
National Agri-Food Innovation Campus
Sand Hutton
York
YO41 1LZ
UK

Tel: +44 (0)1904 462100
Fax: +44 (0)1904 500440
e-mail: info@fapas.com
web: www.fapas.com