ISSN 1757- 5958

amc technical briefs

background paper

Editor: Michael Thompson Analytical Methods Committee AMCTB 18A January 2005

What is proficiency testing?

Guide for end-users of chemical data

These comments are based on misunderstandings reported by users and producers of analytical data. Interpretation of analytical proficiency test results should be, if possible, conducted in collaboration with an analytical chemist.

What is proficiency testing?
Proficiency testing comprises an interlaboratory system for the regular testing of the accuracy that the participant laboratories can achieve. In its usual form, the organisers of the scheme distribute portions of a homogeneous material to each the participants, who analyse the material under typical conditions and report the result to the organisers. The organisers compile the results and inform the participants of the outcome, usually in the form of a score relating to the accuracy of the result.

What is the difference between proficiency testing and accreditation?
Accreditation agencies require analytical laboratories to participate in an appropriate PT scheme where one is available, and demonstrate a system for handling the outcome. This is only one of many requirements of accreditation.

What kinds of materials are distributed?
The materials distributed are as close as possible to the materials being regularly analysed, so that the results of the scheme represent the capability of the laboratories working under routine conditions.

What is proficiency testing for?
The primary purpose of proficiency testing is to help laboratories detect and cure any unacceptably large inaccuracy in their reported results. In other words, it is designed as a self-help system to tell the participants whether they need to modify their procedures. Proficiency tests are not ideally designed for any other purpose, although their results can be used, with due regard to their limitations, and combined with other information, for certain other purposes.

Why are there inaccuracies in analytical results?
All measurement gives rise to inaccuracies, technically known as 'errors' in the measurement community. (The word 'error' here does not imply that a mistake has been made, merely that the outcome of the measurement process varies.) Errors arise because of unavoidable variation in the physical or chemical procedure employed to make the measurement. The measurement of chemical concentration requires far more complicated procedures than typical physical measurements such as length or time. It is straightforward to measure a length to an accuracy of one part in a million, but chemical measurements can seldom be made with an accuracy of better one part in a hundred. Mostly the accuracy is not as good as that, especially if concentrations are very low, for instance as when pesticide residues are being determined in foodstuffs.

AMC Background Papers are designed to provide an interface between analytical scientists and professionals from other disciplines who are involved in the procurement and interpretation of analytical data. Technical issues of current interest are presented.

Is the available accuracy good enough?
That depends on the application. Some analyses have to be extremely accurate. For example, in determining the commercial value of a consignment of scrap gold, the gold content has to be determined with the greatest possible accuracy, better than one part in a thousand – a small error could equate to many thousands of pounds Sterling. In other applications, for example determining the concentration of copper in soil, an accuracy of one part in ten probably suffices – it doesn't matter whether the true value is 20 or 22 ppm if the only decision to be made is whether the level is above or below 200 ppm. Cost comes into consideration as well. As a rule of thumb, to improve the accuracy of a measurement by a factor

 © The Royal Society of Chemistry 2005

of two decreases the chance of an incorrect (i.e., expensive) decision, but increases the cost of analysis by a factor of four. These considerations are known as 'fitness for purpose'.

How do proficiency testing schemes evaluate the accuracy of individual laboratories?

Most schemes convert the participant's result into a 'z-score'. This score reflects two separate features, (a) the actual accuracy achieved (i.e., the difference between the participant's result and the accepted true value), and (b) the scheme organiser's judgement of what degree of accuracy is fit for purpose.

How should z-scores be interpreted?

z-Scores must be interpreted on a statistical (probabilistic) basis and this requires expert knowledge. Here is an outline interpretation.

- A score of zero implies a perfect result. This will happen quite rarely even in perfectly competent laboratories.
- Laboratories complying with the PT scheme's fitness for purpose criterion will commonly produce scores falling between - 2 and 2. They might expect to produce a value somewhat outside this range occasionally by chance, roughly about one time in twenty, so an isolated event of this kind is not of great moment. The sign (i.e., + or -) of the score indicates a negative or positive error respectively.
- A score outside the range from –3 to 3 would be very unusual for a laboratory operating under the given fitness for purpose criterion, and is taken to indicate that the cause of the event should be investigated and remedied.

What mistakes are commonly made in using z-scores?

It is important not to over-interpret z-scores. This could happen in a number of ways.

- Comparing z-scores between rounds or between laboratories has to be done with great caution. A single laboratory operating consistently in line with the fitness for purpose criterion would typically produce z-scores in successive rounds covering the range –2 to +2: the following set [0.6, -0.8, 0.3, 1.7, 0.7, -0.1] would be typical. The small ups and downs between the scores do not indicate a change in performance – they arise by chance. So 1.7 is not 'worse' than 0.3: it does not indicate deterioration in performance.
- Because of this 'natural variation' it is not sensible to make a 'league table' of laboratories based on their z-scores in a round. It is not valid to claim that a laboratory scoring 0.3 in a round is better than another scoring 1.7.
- Judgements based on average z-scores again require caution. Averages of z-scores obtained on a number of different analytes should not be used: they may well hide the fact that one of the analytes consistently gives a poor z-score. Averages of scores from the same analyte over several rounds may be more useful, but still need expert interpretation on a statistical basis.

What are the limitations of proficiency testing?

- Proficiency testing has to be carried out within the context of a complete system for appropriate quality in each laboratory. It cannot be used a substitute for routine internal quality control. It is not a means of training individual analysts, nor (for the participant laboratory) a way of validating analytical methods.
- Proficiency testing provides a participant laboratory only with an indication of problems if they are present. It does not provide any diagnostics to help solve the problem.
- Success in a proficiency test for one analyte does not indicate that a laboratory is equally competent in determining an unrelated analyte.

This Background Paper was produced for the Analytical Methods Committee by the Statistical Subcommittee (Chairman M Thompson),which is supported by the Food Standards Agency. The paper has been endorsed by the Eurachem UK Working Group on Proficiency Testing.

AMC Background Papers, Technical Briefs and Recommendations may be freely reproduced and distributed in exactly the same form as published here, in print or electronic media, without formal permission from the Royal Society of Chemistry. Copies must not be offered for sale and the copyright notice must not be removed or obscured in any way. Any other reuse of this document, in whole or in part, requires permission in advance from the Royal Society of Chemistry. Correspondence should be addressed to: The Secretary, The Analytical Methods Committee, The Royal Society of Chemistry, Burlington House, Piccadilly, London W1J 0BA.

Other AMC products can be found on:
www.rsc.org/lap/rsccom/amc/amc_index.htm