codex alimentarius commission

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00153 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 3a) **CX/MAS 07/28/3**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING
Twenty-eighth Session
Budapest, Hungary, 5 – 9 March 2007

DRAFT GUIDELINES FOR EVALUATING ACCEPTABLE METHODS OF ANALYSIS

(Prepared by New Zealand with the assistance of the Working Group[1])

BACKGROUND

1. Draft Guidelines for Evaluating Methods of Analysis prepared by New Zealand with the assistance of the working group were issued in August 2005 by Circular Letter (CL 2005/44-MAS), and there was a significant discussion on the subject at the 27th session of CCMAS. The committee agreed that further consideration should be given to the general approach and contents of the Draft Guidelines before proceeding to a general discussion. It was agreed that the Delegation of New Zealand, with the assistance of an electronic working group, would redraft the Draft Guidelines taking into account the issues raised in written comments and at the session[2].

GENERAL APPROACH AND CONTENTS

2. Concerning the general approach and contents of the Draft Guidelines, there were several comments that the working group is acting on. Comments suggested that the Guidelines were too complex, leading to possible confusion in interpretation. It was suggested that they should be restructured, simplified, and made more practical and easier to follow. It was also suggested that the new concepts should be considered for publication in a scientific journal before they are considered in the framework of Codex. In order to carry out the work in a logical sequence, the working group decided the first task should be to extract material from the Annexes for publication elsewhere, and to follow this by redrafting the remaining material, taking account of the comments of a technical nature.

3. Removing the more technical material from the paper and publishing it elsewhere will reduce the size and complexity of the document required to specify the recommendations, and the more technical parts can be submitted to peer review in a more suitable forum.

4. It is planned to publish three papers, roughly as follows:

   1) Methods of providing confidence intervals for estimates of precision parameters, particularly the within-laboratory standard deviation, as part of method-performance studies.

   2) Considerations relating to the estimation of bias and its uncertainty of estimation in method-performance studies.

   3) The impact of uncertainty relating to estimates of bias and precision on producer's risk in tests for product compliance, and the incorporation into compliance tests and tests for method acceptability of suitable controls of this possible impact.

5. Progress as of now is that a draft of the first paper has been completed.

---

1 Argentina, Australia, Austria, Brazil, Dominica, European Community, Honduras, Japan, Republic of Korea, Netherlands, South Africa, Spain, United Kingdom, United States of America.

2 ALINORM 06/29/23, para. 22.

6. The Committee's decision to prefer publication in other forums before adopting the recommendations as they stood has proved to be a wise one. The methods initially proposed for the estimation of confidence limits for parameters, which were initially put forward by New Zealand under some pressure of time, although still in common use, turned out on investigation to be poor, and a search for better methods was undertaken. New Zealand has tested several proposed methods by simulation studies, and believes it can suggest methods that are workable in practice, and while not necessarily statistically optimal, give something reasonably close to a correct answer. Further revision is still to be undertaken, which might lead to further changes in the methods.

7. In spite of the slow progress in specifying appropriate statistical procedures, it is still felt worthwhile for the Committee to consider the guidelines. Irrespective of matters of statistical technique, the guidelines put forward an approach to the problem of method acceptability that focuses on the risks of incorrect decisions and the estimation of those characteristics of measurement methods that are needed for the assessment of these risks. The inclusion of detailed prescription of statistical techniques has the disadvantages of substantially increasing the length and complexity of the guidelines and of opening them up to potentially time consuming technical disputes on statistical methodology. It must also be recognized that in many cases modification of recommended statistical methods will be necessary in the light of particular features of a method-performance study. In keeping with the philosophy of fitness for purpose, which applies to statistical methods as well as analytical methods, it is felt that the main purpose of these guidelines should be to express what is to be achieved rather than to prescribe the methods by which it is to be achieved.

8. It is however still considered desirable to provide some advice on statistical methods. In particular, some advice on the estimation of confidence limits for precision parameters may be valuable, as this is outside the range of techniques currently employed in method-performance studies. Such material could be incorporated as additional annexes to the guidelines over the course of time, or included in the Codex recommendations for method-performance studies. It is evident that the time required to complete this work is considerably longer than was envisaged when it was undertaken, and it has not been possible to redraft the Annexes to the Guidelines in time for consideration at this session of CCMAS. The working group considers a realistic target would be the 29th session.

REVISION OF THE CORE GUIDELINES

9. The working group has prepared a revised version of the core Guidelines (i.e. paragraphs 1 – 21, before the Definitions and Annexes) in the light of written comments and comments at the 27th session of CCMAS – see Appendix 1. These amendments all seem reasonably straightforward. Further revision has been made to separate matters of principle from matters of statistical technique. It would therefore be worthwhile, in the interest of making progress, for the committee to consider this material at the current session, with the annexes to follow later.

RECOMMENDATIONS TO CCMAS

Recommendation 1

10. The working group recommends that CCMAS consider the revised core Guidelines (Appendix 1).

Recommendation 2

11. The working group recommends that CCMAS consider the criteria set out for Acceptance of methods set out in the annex.

Recommendation 3

12. The working group recommends that CCMAS agree that the working group should continue work on additional annexes to the guidelines giving recommendations on statistical procedures, and report on progress to the 29th session of CCMAS.

Recommendation 4

13. The working group recommends that Codex guidelines for method-performance studies should be updated to include estimation and reporting of confidence limits for precision parameters, and that recommendations for information required for evaluation of methods submitted for endorsement should be updated.

Appendix 1

DRAFT GUIDELINES FOR EVALUATING ACCEPTABLE METHODS OF ANALYSIS

Scope

1. These guidelines provide a framework for evaluating acceptable methods of analysis, including new methods, modified methods, and methods with a different scope of application.

2. The guidelines apply to methods that may be used for control, inspection or regulatory purposes in relation to import and export of foods.

3. The guidelines specify criteria which methods must satisfy to allow the substitution of one method (presumed fit for purpose) by another. The guidelines specify uncertainty parameters that should be documented to enable an adequate assessment of fitness for purpose to be made.

4. The guidelines will not be applicable in some cases, for example where methods are not available in the public domain or where a method is being developed for a new analyte.

Objectives

5. These guidelines provide a scientific basis for the selection and acceptance of analytical methods to be used in assessments of product in order to protect the consumer and to facilitate fair trade.

6. The guidelines are intended to allow more flexibility, through the development and application of appropriate criteria for methods, through consideration of their fitness for purpose, as the basis for their acceptance.

Requirements

7. The acceptance of a method consists of the steps:

   (a) estimation of the performance characteristics of the method,

   (b) evaluation of the method based on its performance characteristics and its fitness for purpose, and

   (c) formal acceptance of the method.

8. Acceptance of a method may be:

   (a) unconditional, or

   (b) conditional acceptance provided suitable allowance is made to render the method fit for purpose in its intended application.

Estimation of the performance characteristics of a method

9. Laboratories involved in the evaluation should comply with the Codex Guidelines for the Assessment of the Competence of Testing Laboratories Involved in the Import and Export Control of Foods (CAC/GL 27-1997) as amended in the last version of the Guidelines.

10. To the extent possible, the method's characteristics and the error associated with them should be estimated in a method performance study, conducted as recommended in Codex Food Control Laboratory Management: Recommendations (CAC/GL 28-1995, Rev.1-1997 and its updates). An adequate method description and verifiable performance data should be available for peer review.

11. The following performance characteristics of the candidate method should be estimated and assessed as appropriate. [The definitions of these characteristics are in ….]

- bias
- sensitivity.
- linearity
- precision (repeatability, reproducibility, and between-laboratory standard deviation)
- limit[s] of detection[and quantification]
- applicability (analytes, matrix, concentration range and preference given to 'general' methods)
- recovery

- ruggedness (robustness)
- selectivity (interference effects etc.)

12. To enable the assessment of fitness for purpose, at least the following should be reported:

    a) An estimate of method bias over the relevant range

    b) A 95% confidence interval for the method bias

    c) Estimates of the repeatability standard deviation, the between-laboratory standard deviation and the reproducibility standard deviation

    d) Upper confidence limits for the precision parameters in c). If the criteria in the Annex are to be used, the limits given should include 80% limits.

The extent to which current guidelines for method performance studies meet this requirement is discussed in the Note (paragraph 24).

13. In the case of validation by a single laboratory, there is lower confidence in the general applicability of the resulting estimates of statistical parameters than is available from interlaboratory validation. In particular, statistically based estimates of method bias, the between-laboratory standard deviation and the overall reproducibility will not be available, although reasonable guesses at their approximate values may in some circumstances be made based on experience with similar methods. These may be acceptable in certain non-critical applications, for example tests in which the cut-off is so far from the normal range of analyte concentration that the risk of normal product being incorrectly rejected is remote.

14. Some characteristics such as precision and limit of detection can be also be applied in the case of defining methods (Type I).

Evaluation of the method based on its performance characteristics and its fitness for purpose

15. Criteria should be established, involving relevant performance characteristics, for evaluating the acceptability of methods. The criteria may be specified as a requirement.

16. Criteria should take account of:

- The performance characteristics of existing accepted methods; or
- Fitness for purpose considerations in relation to the intended use of the results to evaluate conformity. These will normally include consideration of the potential impact of the measurement uncertainty associated with the method on the risks of incorrectly accepting or rejecting product at various analyte levels. Criteria for using this approach are given in the Annex.

Commentary

The Harmonised Guidelines for Internal Quality Control in Analytical Laboratories define Fitness for Purpose as the "degree to which data produced by a measurement process enables a user to make technically and administratively correct decisions for a stated purpose".

In terms of Codex, the purpose will generally be assessments of conformity against Codex standards. In this context Fitness for Purpose of a method is interpreted to mean that measurement error and bias of the method will not cause unacceptable increases in the producers' and consumers' risks (or equivalently the AQL and LQ) for a particular assessment of conformity. In addition, given that estimates of bias and precision parameters obtained from method-performance studies may be subject to considerable uncertainty, uncertainties of these estimates must also be taken into account when assessing fitness for purpose.

17. Assessment of fitness-for-purpose criteria may suggest a need to alter the method of evaluating conformity, for example by specifying tolerances or adjusting existing tolerances to allow for measurement uncertainty or by changing the number of samples used in the test.

18. In assessment of fitness for purpose, it should be recognized that estimates of bias and precision parameters from a method-performance study may themselves be subject to appreciable statistical uncertainty, and if possible an attempt should me made to allow for this by using conservative estimates, for example, those based on upper confidence limits for the precision parameters if these are available.

19. A candidate method may be accepted for general use if it satisfies the conditions outlined in the Annex.

Formal acceptance of the method

20. Methods may be formally accepted by a country by:

- Approving specific methods that satisfy criteria, or
- Identifying or developing methods that satisfy specified criteria (see 15).

21. When a method is accepted, its description, estimates of the relevant characteristics, the criteria that were used in assessing the method and the demonstration that the criteria have been met should be formally documented.

22. The scope of the acceptance should be determined according to the range of experimental conditions under which the method has been tested. For example, in cases where performance characteristics have been estimated in fewer laboratories than specified in the recommendations for method-performance studies, the acceptance may be restricted.

23. If a method of analysis has been endorsed by Codex, then preference should be given to using that procedure.

Note

24. The recommendations given in paragraph 12 extend the current recommendations for method-performance studies. For example, they extend current IUAPC guidelines in at least the following respects:

a) In the guidelines, estimates of method bias are not explicitly required. The mean for each sample is to be reported along with the "true or accepted value, if known." The qualification "if known" would seem to imply that the estimation of bias from a method performance study is not strictly necessary. However, the question of method bias does need to be addressed, either by use of reference material, by parallel analysis using a standard method, or by other means. The absence of a statistically significant bias does not of itself imply that the bias is negligible, but only that it has not been precisely enough estimated to be sure of its existence or direction. Consequently a confidence interval should be given so that its possible values may be allowed for in assessing fitness for purpose.

b) In the guidelines, reporting of the between-laboratory standard deviation is not required. This precision parameter is important in assessing the measurement uncertainty relating to a mean of several samples. In the absence of this parameter the precision of such a mean would have to be conservatively estimated by the overall reproducibility. Although the between laboratories standard deviation can be estimated directly from the repeatability and reproducibility standard deviations, which it is required to report, its confidence limits cannot be deduced from theirs.

c) Confidence limits for precision parameters are not normally required. This may be partly due to the difficulties inherent in calculating them, particularly the between laboratories standard deviation. Lack of balance in the trial design to missing data or rejection of outliers can also cause considerable difficulty. However, considerable research into this problem has taken place, and reasonable methods have been reported in the literature and are finding their way into statistical packages, to replace the traditional "Satterthwaite approximations," which perform very poorly in some contexts. The size of these confidence limits when they are calculated sufficiently demonstrates the need for them.

ANNEX: CONDITIONS FOR GENERAL ACCEPTANCE OF METHODS

This Annex gives the conditions under which a candidate method may be accepted as a general replacement for a standard method in judging product conformance after a method-performance study.

Failure to meet the conditions does not exclude a method from use in particular circumstances where it is judged "fit for purpose." However, it is envisaged that only exceptional circumstances would justify objections to the use of a candidate method that does meet the conditions.

The conditions on bias and precision are formulated to satisfy the following criteria:

CRITERIA

a) Consider a test of a single sample or mean of several samples against a cut-off. Suppose that uniform product with a certain analyte concentration has a 5% probability of failure when the standard method is used, according to the accepted precision parameters of the standard method.

Then the probability of failure of this same product under the candidate method should not exceed 7.5%.

b) The estimates of method parameters should be of sufficient precision to guarantee that condition a) above is met with at least 80% confidence.

NOTES

1. In the absence of method bias condition a) is met provided that neither the repeatability standard deviation nor the between laboratories standard deviation increases by more than 14%.

2. The size of trial required to confidently rule out method bias of a size that would of itself violate condition a) would normally be considered excessive. Thus an explicit bias correction is required in all cases.

3. In condition b) the level of confidence sought seems at first sight rather low. However, to increase it substantially would make it very difficult for a method of the same precision as the standard method to be successfully validated in a trial of moderate size.

CONDITIONS

Trial design

To plan and carry out a method-performance study is a major undertaking involving scientific and statistical considerations that will vary considerably according to the analyte and method being considered. No prescription is therefore made on the type of investigation or the methods of estimation that should be used, other than that they should yield scientifically and statistically valid estimates or assessments of the performance characteristics listed below. It is however thought unlikely that satisfactory estimates of bias and precision will be obtained unless a substantial inter-laboratory trial is included. This would normally be conducted according to the current Codex recommendations for method-performance studies.

Bias

Upper and lower one-sided 95% confidence limits for method bias (or relative method bias, if this is more appropriate) at various concentrations within the range should be given, for the purpose of calculating appropriate adjustments when product is tested against a specification limit. (Note that together these limits will form a 95% confidence interval.)

Applicability

The range of matrices and concentrations for which the method has been tested and found appropriate should be given. Known matrices for which the method is unsatisfactory should be reported.

Limit of detection[, limit of quantification]

Should be reported.

Linearity

Statistically significant evidence of non-linearity should be reported, and the confidence bounds for method bias must be calculated in a way that allows for the non-linearity. This would normally be done by calculating independent estimates of bias at various parts of the range.

Precision

An upper 80% confidence bound should be calculated for the repeatability standard deviation. This should not exceed the accepted value for the repeatability standard deviation of the standard method by more than 14%.

An upper 80% confidence bound should be calculated for the between laboratories standard deviation ($\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2}$). This should not exceed the accepted value for the corresponding standard deviation of the standard method by more than 14%.

An upper 80% bound for the reproducibility standard deviation should also be given.

Failure to meet these requirements would require a tolerance to allow for the possible additional measurement uncertainty.

Recovery

Recovery rate, if not explicitly corrected for as part of the methods will form an element of the method bias, and accordingly no additional conditions are required. Anomalous recovery rates for particular matrices should be reported, and the matrices concerned excluded from the range of applicability of the method.

Selectivity (interference effects etc.)

Any interference effects should be reported.

Sensitivity

No conditions are placed on sensitivity, partly due to the difficulty of estimating it in certain circumstances. Estimates of sensitivity when these can be made will normally be incorporated into the confidence intervals for bias.

NOTES

Failure of a method to meet the precision criteria means that an additional tolerance needs to be applied in compliance testing when the method is used, to allow for the possible additional measurement uncertainty. This would be applied on top of the adjustment for method bias. There is a possibility that increased measurement uncertainty could be compensated for (from the producer's point of view) by, for example, a negative upper 95% bound for bias when testing against an upper limit. This could render the method fit for purpose in some contexts. However, it is not currently proposed that such a method should be considered "generally acceptable", as the method would on the face of it appear likely to be inferior, with a statistically significant bias and at least the possibility of substantially poorer precision than the standard method.