codex alimentarius commission

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00100 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 3a) **CX/MAS 06/27/3**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING
Twenty-seventh Session
Budapest, Hungary, 15-19 May 2006

DRAFT GUIDELINES FOR EVALUATING ACCEPTABLE METHODS OF ANALYSIS GOVERNMENT COMMENTS AT STEP 6 (Argentina, Australia, Brazil, Cuba, Hungary, Japan, Switzerland)

ARGENTINA (English version)

REFERENCES

Struck out text = Original text in document sent out for comments
Double struck out text = Text deletion proposed by Argentina

Underlined, black text = New text proposed by Argentina

Scope

1. These guidelines provide a framework for evaluating acceptable methods of analysis.

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 be used as Type III methods. Some of the considerations may also apply to defining methods (Type I).

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

Objectives

5. These guidelines ~~are intended to assist countries in the application of requirements for trade in foodstuffs~~ 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 as the basis for their acceptance.

Requirements

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

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

(b) ~~judgement~~ **Evaluation** of the method based on its performance characteristics and its ~~fitness~~ **appropriateness** for purpose, and

(c) ~~formalising~~ **Formal** acceptance of the method.

Estimation of the performance characteristics of a method

8. 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**.

9. The following performance characteristics of the candidate method should be estimated ~~assessed as appropriate against the following criteria by laboratories involved in the import and export control of foods~~:

- ~~Accuracy~~
- bias
- sensitivity.
- linearity
- precision (repeatability, reproducibility, and reproducibility net of repeatability)
- limit[s] of detection [and quantification]
- applicability (analytes, matrix, concentration range and preference given to 'general' methods)
- recovery
- ruggedness (robustness)
- selectivity (interference effects, etc.)
- **uncertainty**

10. [Definitions of these characteristics are given below, and the method for the corresponding estimations ….]

11. 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**). The data from the method performance study should be analysed as described in Annex A. An adequate method description and verifiable performance data should be available for peer review.

12. In the case of ~~single-laboratory~~ validation **by a single laboratory**, there is lower confidence in the general applicability of the resulting estimates of statistical parameters than is available from studies of ~~several-laboratory~~ **interlaboratory validation**.

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

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

14. **For ~~judging~~ the acceptability of methods** criteria should be established, involving relevant performance **evaluation** characteristics~~, for judging the acceptability of methods~~. The criteria may be specified as a requirement.

15. Criteria should take account of:

- The performance characteristics of existing accepted methods. An example of this approach is given in Annex C;
- Importing country's requirements or requirements by the consumer;
- Specifications in food standards; or
- ~~Fitness~~ **Appropriateness**-for-purpose considerations in relation to the intended use of the results to judge conformity.

16. Assessment of **appropriateness** ~~fitness~~-for-purpose criteria may suggest a need to alter the method of **evaluating** ~~judging~~ conformity.

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

~~Formalising~~ Formal acceptance of the method

18. Methods may be formally accepted ~~in the~~ **by a** country by:

- Approving specific methods that satisfy criteria, or
- Identifying or **developing** methods that satisfy specified criteria (see 14).

19. When a method **is accepted**, ~~the~~ **its** description ~~of the method~~, estimates of the relevant characteristics and the demonstration that the criteria have been met should be formally documented.

20. The scope of the acceptance should be determined ~~by~~ **according to** the range of experimental conditions under which the method has been tested. For example, **the acceptance may be restricted** in cases where performance characteristics have been estimated in fewer laboratories than specified in Annex A.

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

NOTE:
All definitions are proposed to be included in a single, easy-to-update document or as annex to an existing one.

ARGENTINA (versión en español)

Referencias
Texto TACHADO = ORIGINAL EN EL DOCUMENTO ENVIADO PARA OBSERVACIONES
Texto TACHADO DOBLE = ELIMINACIÓN DE TEXTO SUGERIDA POR ARGENTINA
Texto en NEGRITA y SUBRAYADO = NUEVO TEXTO SUGERIDO POR ARGENTINA

Ámbito de aplicación
1. Estas directrices proporcionan un marco para la evaluación de métodos de análisis aceptables.
2. Las directrices se aplican a los métodos que pueden utilizarse para el control, inspección o fines reglamentarios en relación con la importación y exportación de alimentos.
3. Las directrices especifican los criterios que deben satisfacer los métodos para utilizarse como métodos del Tipo III. Algunas de las consideraciones pueden aplicarse también a la definición de métodos (Tipo I).
4. Las directrices no serán aplicables en algunos casos, por ejemplo, cuando ~~no se dispone de métodos en el sector~~ **cuando los métodos no son de dominio público** o cuando se está elaborando el método para un nuevo analito.

Objetivos
5. Estas directrices ~~tienen por finalidad ayudar a los países a aplicar los Requisitos para el comercio de productos alimenticios~~ proporcionan una base científica para la selección y aceptación de métodos analíticos que han de usarse en evaluaciones de productos con miras a proteger a los consumidores y facilitar un comercio leal.
6. Las directrices tienen por objeto permitir una mayor flexibilidad ~~mediante la elaboración~~ **a través del desarrollo y aplicación** de criterios apropiados para los métodos, como base para su aceptación.

Requisitos
7. **La** aceptación de un método **consiste en los siguientes pasos**:
a) estimación de las características de rendimiento del método,
b) ~~juicio~~ **Evaluación** del método basado en sus características de rendimiento y en su ~~aptitud para la finalidad~~, **adecuación al propósito** y
c) ~~formalización de la aceptación~~ **Aceptación formal del método**.

Estimación de las características de rendimiento de un método
8. Los laboratorios que participen en la evaluación deberían ajustarse a las Directrices del CODEX para la evaluación de la competencia de los laboratorios de ensayo que participan en el control de las importaciones y exportaciones de alimentos (CAC/GL 27-1997) **con adecuación a la última versión de la Norma.**
9. ~~Los laboratorios que participan en el control de las importaciones y exportaciones de alimentos deberían evaluar~~ Se deberían estimar las siguientes características de rendimiento del método propuesto:

- ~~Exactitud~~
- sesgo
- sensibilidad.
- linealidad
- precisión (repetibilidad, reproducibilidad, y reproducibilidad menos la repetibilidad)
- límite[s] de detección [y cuantificación]
- aplicabilidad (analitos, matriz, intervalo de concentración y preferencia dada a los métodos "generales")
- recuperación
- solidez (robustez)
- selectividad (efectos de interferencia, etc.)
- **Incertidumbre**

10. [Las definiciones de estas características se presentan a continuación y el método para las estimaciones correspondientes….]
11. En la medida de lo posible, las características del método y el error asociado con ellas deberán estimarse en un estudio del rendimiento del método, realizado según lo recomendado en Gestión de Laboratorios de Control de Alimentos: Recomendaciones (CAC/GL 28-1995, Rev.1-1997 **y sus actualizaciones**) Los datos del estudio del rendimiento del método deberán analizarse según se describe en el Anexo A. Deberá

facilitarse para un examen por especialistas una descripción adecuada del método y datos verificables del rendimiento.
12. En caso de una ~~única validación de~~ **validación por un único laboratorio**, la confianza en la aplicabilidad general de las estimaciones resultantes de parámetros estadísticos es menor que si se dispone de estudios de ~~varios laboratorios~~ **validación por interlaboratorios**.
13. Algunas características, como la precisión y el límite de detección, pueden aplicarse también en el caso de métodos de definición (Tipo I).

~~Juicio~~ Evaluación del método basado en las características de rendimiento y en su adecuación al propósito ~~aptitud para la finalidad~~

14. **Para ~~juzgar~~ la aceptabilidad de los métodos** deberían establecerse criterios, que incluyan características pertinentes de **la evaluación del** rendimiento, ~~para juzgar la aceptabilidad de los métodos~~. Los criterios podrían especificarse como un requisito.
15. Los criterios deberían tener en cuenta:

- Las características de rendimiento de métodos aceptados existentes. En el Anexo C se ofrece un ejemplo de este enfoque;
- Los Requisitos del país importador o los Requisitos del consumidor;
- Especificaciones incluidas en normas alimentarias; o
- Consideraciones de ~~aptitud para la finalidad~~ **adecuación al propósito** en relación con el uso a que se destinan los resultados para juzgar la conformidad.

16. La evaluación de los criterios de ~~aptitud para la finalidad~~ **adecuación al propósito** puede sugerir la necesidad de modificar el método ~~de juicio~~ **para evaluar** ~~de~~ la conformidad.
17. Puede aceptarse un método propuesto para uso general si cumple las condiciones descritas en el Anexo B.

~~Formalización de la~~ Aceptación formal del método

18. Los métodos pueden ser aceptados formalmente ~~en el país~~ **por un país** mediante:

- La aprobación de métodos específicos que cumplan **con** los criterios, o
- La identificación o ~~elaboración~~ **desarrollo** de métodos que cumplan los criterios especificados (véase 14).

19. Cuando **se acepta** un método ~~un método~~, deberá~~n~~ documentarse formalmente la descripción del ~~método~~ **mismo**, las estimaciones de las características pertinentes y la demostración de que se han cumplido los criterios.
20. El Ámbito de aplicación de la aceptación deberá determinarse ~~con arreglo~~ **de acuerdo** a las distintas condiciones experimentales en que se ha ensayado el método. Por ejemplo, **la aceptación podrá ser restringida**, en casos en los que las características de rendimiento se han estimado en un número menor de laboratorios que el especificado en el Anexo A.
21. Si un método de análisis ha sido ratificado por el CODEX, deberá darse preferencia a la utilización de ese procedimiento.
NOTA:
Se propone que todas las definiciones estén contenidas en un único documento de rápida actualización o en anexo a un documento ya existente.

AUSTRALIA

General Comments

Australia supports the need for guidelines for evaluating methods of analysis, particularly given the shift by CCMAS to the criteria approach for the endorsement of methods of analysis. However, significant improvements to the document are still required, in order for it to be a useful tool for the Codex method endorsement process. Australia's primary concern is that the draft guidelines are overly complex and may create confusion in interpretation. Furthermore the paper in its current form does not address situations for single laboratory validated methods in a balanced manner. These methods may be considered "acceptable methods of analysis" by other horizontal Codex Committees. Given the terms of reference of CCMAS it is

essential that clarification is given either on the particular methods of analysis under consideration or situations which include methods of analysis relevant to all Codex committees.

It is particularly important that terminology, including parameters for equations are clearly and consistently defined throughout the document and that the various sections and annexes of the document are complementary. As it stands a number of new and different issues and terminologies are raised throughout the various parts of the document, which do not effectively aid in the understanding of significant principles. Some of these issues could be improved by a more comprehensive use of literature references throughout the document. The use of "error" as opposed to "uncertainty" throughout the document should be reviewed in order to ensure that the correct meaning is applied, that is, consideration to the use of the term "uncertainty" should be given, where appropriate. Annex A raises a number of new issues not covered in the text, which are also not suitably defined and can lead to further interpretative difficulty.

Australia would suggest that the drafting group restructure the document into a simplified version, in order to assist with the flow of information and appreciation of the principles for evaluating acceptable methods by all member countries and Codex committees.

Specific Comments

REQUIREMENTS

Estimation of the performance characteristics of a method

9. Australia suggests reference to "reproducibility net of repeatability" is not common terminology and should be supported by a more suitable explanation in Annex A, in order to avoid confusion.

Formalising acceptance of the method

21. "If a method of analysis has been endorsed by Codex, then preference should be given to using procedures endorsed by Codex".

Australia seeks clarification on the intent of this statement. It is unclear whether member countries or Codex committees will give preference to the use of methods - whether "use" is for the purposes of prescription to a numerical standard or for the purposes of application of a methodology for a particular analyte-food combination.

[DEFINITIONS

Repeatability [Reproducibility]

Australia suggests "Precision under repeatability [reproducibility] conditions" be modified to "Precision **of a method** under **stated** repeatability [reproducibility] conditions" for clarity.

Within-laboratory reproducibility standard deviation

Note 2 under this definition, requires modification in order to provide useful clarification on "within-laboratory reproducibility standard deviation".

Standard deviation for reproducibility net of repeatability

The parameter σ_L requires a clearer definition and further clarification, particularly with respect to the various functions within the equation.

"The standard deviation $\sigma_L = \sigma^2_R - \sigma^2_r$ where σ_R is the **between laboratory (i.e. variance between laboratories)** reproducibility standard deviation and σ_r is the **within laboratory (i.e. variance of one laboratory)** repeatability standard deviation.

Australia suggests "measurement **uncertainty**" rather than "measurement **error**" be used in this context, as indicated in the general comments above. "Error" refers to bias from the true value whereas "uncertainty" defines the range of the values that could reasonably be attributed to the measured quantity.

ANNEX A: ESTIMATION OF CHARACTERISTICS

BIAS

The terminology "error of indication" and "expectation" is not common terminology thus requires further explanation in order to assist the reader's understanding and should either be more comprehensively addressed within the document or referenced appropriately.

General Comments on Bias

As indicated in the general comments a clear definition of each of the parameters for the equation $b(x) = b(x_0) + (s-1)(x-x_0)$ is required.

Some indication of how the parameter "*s*" - the estimated sensitivity is calculated must be provided in order for this equation to be of relevance to the general comments on bias, either in this section, in the context of estimation of bias or under the section on *Estimation of Sensitivity*. Without this, the entire discussion on bias appears to be somewhat redundant.

Reference to "standard error" should be modified to "standard uncertainty" given that this is not a measurement of the difference from the true value as implied by the term "error".

SENSITIVITY

Estimation of sensitivity

Australia offers the following suggested modifications in order to assist with clarity of understanding under this section:

"3. The standard deviation of 2) is divided by the square root of the number of laboratories to give an estimate of the standard error **[uncertainty]** of the **mean** estimated method sensitivity."

"For comparing the range of reference values with their uncertainties, it is difficult to give criteria for smallness **of the uncertainty in the reference values** that are generally valid."

".... where σ_e is the standard deviation of the errors in the reference values **of the reference materials** and σ_x the standard deviation of the reference values themselves...." Please note: It is unclear how the reference values are derived – some clarification whether this is through reference materials or reference methods is required.

Clarification of the parameters β, σ_e and σ_x is required. Particularly useful would be some explanation of how these parameters are derived mathematically.

LINEARITY

The definition of linearity does not quite clarify whether linearity in this context refers to linearity of bias or linearity of calibration.

Modification of the sentence: "With the minimum number of five samples....patterns in the reference values (eg. predominantly convex shape) are present...." to "With the minimum number of five samples.... patterns in the reference values (eg. **normally distributed**) are present...."

Although the guidelines advise that in the case of non-linearity this solution is not available they provide little guidance on cases of non-linearity. Clear guidance on these situations may be more useful. The observation that "The only solution may seem to be to considerably increase the number of samples used within the concentration range" falls short of being instructive in line with the established intent of useful Codex guideline documents.

PRECISION (REPEATABILITY, REPRODUCIBILITY, REPRODUCIBILITY NET OF REPEATABILITY)

Australia suggests that Note 1 is modified from "Precision depends only on the distribution of random errors and does not relate to the true value or the specified value" to "Precision depends on the **random** distribution of **results** and does not relate to the true value or the specified value" for the purposes of clarity.

It is unclear whether "specified value" refers to "reference value". If so, "specified value" should be replaced with "reference value" for consistency. If "specified value" refers to a different value some indication of how it is derived would be required.

Note 3, Australia suggests the word "Quantitative" is deleted to read "Measures of precision depend critically on the stipulated conditions."

Estimation

See comments under Annex B regarding reference to "an upper 80% confidence interval".

Australia suggests "The statistical model involved is" is modified to "The **model for the measured concentration of a sample is**".

The purpose and how the model is derived remain unclear.

The equation provided in this section for σ_L is inconsistent with definitions provided under the DEFINITIONS section (pg 7) and the EXPLANATORY NOTES FOR ANNEX B section (pg 30).

1. Obtaining the estimates and estimates of their sampling variances
2. Obtaining the confidence intervals

Although these sections attempt to clarify for the reader the two steps involved in obtaining confidence limits for the standard deviations, there are a number of parameters that are not sufficiently defined or proven. Literature references that support the content and provide further background on particular factors, for example "Satterthwaite's Approximation" are recommended.

LIMIT OF DETECTION

Estimation

The percentage quoted in the following sentence "This estimation corresponds, at least nominally, to a 0.1% chance of declaring the analyte present when the result truly arises from a blank sample." seems too low for an estimate based on 3s. Australia suggests that the percentage "0.1%" be modified to "**1%**".

LIMIT OF QUANTIFICATION

The general comments on Limit of Quantification do not provide sufficient guidance. Australia suggests that Codex guidelines should clearly specify an approach agreed to be the best approach, once the conventional ways for determining LOQ have been assessed, in order to ensure comparability of methods. At the very least the guidelines should ensure that the method specifies how LOQ was calculated.

EXPLANATORY NOTES FOR ANNEX A: ESTIMATION OF CHARACTERISTICS

The explanatory notes for Annex A appear to be a mixture of additional information and definitions rather than "explanations". This section may require modification to make it more concise and clear.

ANNEX B: CONDITIONS FOR ACCEPTANCE OF METHODS

Annex B provides excellent context for the guideline document and is presented in a reasonably succinct manner. Given that Annex B provides the rationale for the guideline document, it would seem appropriate for the information in Annex B to be provided upfront i.e. before what is currently Annex A. Australia notes further work on ensuring consistency in terminology with the previous sections of the document is still required.

RATIONALE

It is unclear whether reference to the use of "the standard method" under this section is the same as the use of "the **reference** method". If the "standard method" method refers to the "reference method" Australia suggests the wording is modified for consistency, otherwise further explanation of "standard method" is required.

CONDITIONS

Bias

It is unclear from Annex A how to calculate the "upper and lower one-sided 95% confidence limits for method bias".

Precision

The percentages (i.e. 80% and 14%) provided for calculating various parameters in this section appear to be somewhat arbitrary. The basis for these numbers is not sufficiently explained elsewhere in the document. Australia suggests that the use of literature references may be applicable in this situation.

EXPLANATORY NOTES FOR ANNEX B: CONDITIONS FOR ACCEPTANCE OF METHODS

It is unclear how these explanatory notes are related to the content of Annex B.

ANNEX C: EXAMPLES OF USE OF METHODS DESCRIBED IN ANNEX A

This Annex serves to clarify the approach defined in Annex A and proves to be very useful, however it would be useful to have more raw data provided, in order for the reader to ensure his/her understanding of how to repeat various calculations.

BRAZIL

A – It doesn't approach scope, objectives, and application conditions in a clear way

Suggestions:

Item 2 to substitute "used for control inspection or regulatory purposes in relation to import and export of foods" by "for evaluating acceptable methods of analysis"; and to include after text: "new methods, modified methods, methods with different application"

Item 5 to exclude the text "in order….far trade" and to include: "for Codex purpose"

Including after of the word Requirement: "Laboratory must attend Requirement from ISO/IEC 17025"

Item 7 Including: Analytical requirement elements and related performance characteristics for example table I EURACHEM GUIDE. The Fitness for Purpose of Analytical Methods. A Laboratory Guide to Validation and Related Topics. Dec. 1998

B – Separated item (definition, obtaining, acceptance conditions and examples) turns the use difficult and compromise the understanding.

We suggest putting together each characteristic performance and its definition, estimation, conditions for acceptance and examples.

C – It doesn't adopt the same established criteria used in previously approved documents by Codex (single, method-performance, quality control, etc).

We suggest to adopt the endorsed criteria by Codex (single, method-performance, quality control,etc: THOMPSON, M; ELLISON, S.L.R. & WOOD, R. Harmonized guidelines for single-laboratory validation of methods of analysis. Pure Applied Chemistry, 74(5): 835-55, 2002; HORWITZ, W. Protocol for design, conduct and interpretation of method-performance studies. Pure & Applied Chem. v.67, nº 2, pgs.331-343, 1995; ISO/IUPAC/AOAC. Harmonized guidelines for internal quality control in analytical chemistry laboratories. Pure & Applied Chemistry. v.67, nº4, p.649-666, 1995; THOMPSON, M.; ELLISON, S.R.L.; et al. IUPAC-ISO-AOAC INTERNATIONAL. Harmonized Guidelines for the Use of Recovery Information in Analytical Measurement. Pure Applied Chemistry, 71(2): 337-348, 1999;

Utilizing established bibliography by recognized organizations about validation, for example: Eurachem, EPA, Codex (criteria), FDA, AOAC.

D – It presents statistical analysis that is of difficult application by laboratories, without justification from statistical procedure adoption and without functionality from purpose statistical test.

Example: Some considerations about purpose linearity test.

Experimental design

It has just purposed various experimental designs.

CL 2005/44 recommends a minimum of 5 levels, in duplicate for evaluating of the non-linearity, quoting the IUPAC (Danzer & Kurrie, 1998). However, Thompson et al., (2002) recommend 6 or more calibration standards spaced equally in interest range, preferentially triplicate or more.

Opinion: The number of liberty degree for a design with 5 levels, in duplicate is so small for a evaluating of the non-linearity.

Recommended test

The recommended method by CL 2005/44 for testing the non-linearity was also quoted by IUPAC (Danzer & Kurrie, 1998), among various other tests for evaluating the linearity. In reality, this test was purposed to Mandel (1964), shown Mandel's fitting test.

This test compares 2 possible fitting models, a linear and other polynomial (it was recommended the cubic by CL 2005/44), comparing the obtained variances by means F test.

$$F = \frac{\left(SQ_{Linear} - SQ_{Curva}\right) GLR_{Curva}}{SQ_{Curva}\left(GLR_{Linear} - GLR_{Curva}\right)}$$

SQ: square sum GLR: residual' liberty degree

Danzer & Kurrie, (1998) even quote 1) visual inspection residues' graphic; 2) F lack-of-fit for testing the linearity 3) homoscedasticity testing by Fmax test or Bartlett 4) t test for evaluating the parameters of calibration; does also reference the premise of normality. Among the recommendations from Thompson et al. (2002) quote 1) residue evaluation 2) F test lack of adjustment, against pure error (lack-of-fit); heteroscedasticity verification; 3) not use of correlation coefficient.

Suggestion (*): Design: 1) at least 6 levels of concentration; 2) spaced equally; 3) at least 3 replicates "real and independent" of each level; 4) zero level for adjusting the instrument zero; 5) random order of reading.

Analysis: 1) visual inspection of residue graphic; 2) outliers verification ("Jacknife" residue); 3) evaluation of residue normality (Ryan-Joiner); 4) homoscedasticity testing (Durbin-Watson); 6) adjustment to model verification (ANOVA Lack-of-fit)

2) About recommended statistical test

A single F test was purposed to complex problem;

We don't know works that evaluate the power this test. The test mustn't be potent to a little number of points;

Other premises that might be verified to apply the regression method (normality, homoscedasticity, independence)

Final opinions

1) Design seems faulty: there are little points

It doesn't detail the space between the levels, that effects the "leverage" (a little distant can give effect lever)

It doesn't refer independence reading (single repetitions of reading effect the lack-of-fit).

-2) About recommended statistical test

A single F test was purposed to complex problem.

We don't know works that evaluate the power this test. The test mustn't be potent to a little number of points.

Other premises aren't evaluated that might be verified to apply the regression model (normality, homoscedasticity, independence)

(*) Souza SVC, Junqueira RG. A procedure to assess linearity by ordinary least squares method. Analytica Chimica Acta, 552 (2005) 25–35.

Opinion: It is little likely that adjustment Mandel test has enough power for testing the linearity, as a single testand utilizing little liberty degree (5 levels in duplicate).

E – Defining minimum necessary information to the report of validation: results' presentation; conclusion.

F – Quoting the references from utilized bibliography .

CUBA

The first version of the document in English was difficult to analyse and did not facilitate the preparation of criteria in time due to the eminently technical nature of the document, that uses a very specialized terminology and introduces new definitions. Its reading and analysis would have been facilitated it the Spanish version had been available in time.

Our comments are as follows:

1. The document is still very extensive and explicit in the concepts and theoretical assumptions, and although they are very rigorous and scientific, they do not bring the dynamism expected in its practical application as a tool to evaluate acceptable methods. This sometimes makes it difficult to follow the main idea that the document intends to convey. Some formula are presented without clear previous explanation and therefore are not easily understandable.
2. With a view to a synthesis of the final document, the definitions should be omitted and reference should be made only to those mentioned in other Codex documents.

3. The new definition of “Standard deviation for reproducibility net of repeatability” is mentioned in the Definitions, however in Annex B, under the title “Precision”, 2nd paragraph, “Standard deviation for repeatability net of reproducibility” has an identical definition, and this should be harmonized.

CUBA (versión en español)

La primera versión del documento en idioma inglés dificultó su análisis y no favoreció la emisión de criterios en el tiempo establecido por ser un documento eminentemente técnico, que utiliza una terminología muy especializada e introduce nuevas definiciones. Su lectura y análisis habría logrado resultados mucho más útiles de haber contado a tiempo con la versión en español.

Nuestras observaciones son las siguientes:

1. El documento es aún muy extenso y explícito en conceptos y suposiciones teóricas, que si bien son muy rigurosas y científicas, no aportan el dinamismo esperado para su utilización práctica como herramienta de trabajo para evaluar los métodos aceptables. En ocasiones, esto puede dificultar la captación de la idea central que se desea transmitir. Se presentan formulaciones cuya explicación previa no es evidente y por tanto no son fácilmente comprensibles.
2. Con el objetivo de sintetizar el documento final deberían omitirse las definiciones y sólo hacer referencias a ellas mencionando los otros documentos del Codex donde aparezcan.
3. Se menciona en Definiciones la nueva definición “Standard deviation for reproducibility net of repeatability”, sin embargo en el Anexo B, bajo el título de Precisión, 2° párrafo, se define con idéntica fórmula la definición “Standard deviation for repeatability net of reproducibility”, lo cual debe ser armonizado."

HUNGARY

Scope, Objectives and Requirements are acceptable as worked out by the WG.

Comments to Annex A: Estimation of Characteristics

It is good for the reader that the definitions of characteristics and their estimation are given in the same para. We suggest to give definitions rather from VIM than ISO Draft 3234, VIM is more connected to measurements. Note the current work of Inter-Agency Meeting „Harmonisation of Analytical Terminology in Accordance with International Standards”

ad 3 Reference values: reference values of samples – even artificials – always have uncertainty. No need to emphasise here the difficulties originated from uncertainty, it will be exposed later in Explanatory Notes for Annex A

ad 5 Sensitivity: on p.13 Often true concentration will be unavailable... Correctly: While true concentration is not available....

Comments to Annex B: Conditions for Acceptance of Methods

The chapter is clear, short, acceptable.

ad Explanatory Notes for Annex B: It is really an extensive statistical discussion of acceptance conditions. Too much attention is paid to sampling problem, it would be better to separate from the assessment of a new method.

Summary

The Guideline gives the possible estimation methods for performance characteristics of candidates Codex methods, but these should be considered as suggestions rather than prescriptions. As written in the Draft Guidelines (Requirements point 11) : „The data from the method performance study should be analysed as described in Annex A. An adequate method description and verifiable performance data should be available for peer review.”

Thus, the peer review will decide on the acceptance. This Guideline is an important and useful help both in the assessment and at the acceptance of a candidate method.

An extensive list of bibliography with reference numbers is necessary at the end.

JAPAN

Japan would like to thank New Zealand for preparing the Draft Guidelines for Evaluating Acceptable Methods of Analysis.

We are pleased to submit the following comments. We apologize for the delay in sending you these comments.

Draft Guidelines for Evaluating Acceptable Methods of Analysis

Para 9, 6th bullet: Limits of detection and quantification

"Limit of Quantification" should be retained as one of method performance characteristics. The information on limit of quantification is necessary for deciding whether an analytical method is suitable at and around a value required by regulation.

The term "limit of quantification" is used in the draft text of the Guidelines as well as in the draft revision of the analytical terminology for Codex use whereas the terms "limit of determination" and "determination limit" are used in the Procedural Manual as well as in the *Recommendations for a Checklist of Information Required to Evaluate Methods of Analysis and Sampling for Endorsement*. We propose to use consistently only one of these terms throughout Codex in order to avoid any confusion.

Reference Sample

The term "reference sample(s)" is used in Annex A, Explanatory Notes for Annex A, and Annex C. We propose to replace the term "reference sample(s)" with "reference material(s)" which has already been defined in ISO Guide 30.

Annex A: Estimation of Characteristics

3 Reference Values

We propose to amend the second bullet in "3 Reference Values". Proposed insertion and amendments are underlined. We consider the term "difficulties" is more appropriate than "problems" in this case.

"2. Reference ~~samples~~ materials could be used. If possible, certified reference materials (CRM) should be used. In this and the following case the uncertainty attached to the reference values could create ~~problems~~ difficulties in estimating some performance parameters with sufficient accuracy."

SWITZERLAND

Switzerland welcomes the opportunity to submit comments on the "Draft Guidelines for Evaluating Acceptable Methods of Analysis" at step 6 of the Codex Step procedure.

General comments:

We believe that standardization of the criteria to evaluate acceptable methods of analysis at a global level would be helpful and might increase food safety and facilitate fair trade. We also admit that this document has certainly brought the discussion to statistically correct grounds.

Nevertheless, it could be feared that the proposed approach might be relatively far from common practice in most control laboratories. As stated on page 1 of Annex C, in a different context, "its use requires a statistically skilled analyst".

The aspects of fitness-for-purpose are not sufficiently taken into account, e.g. "Limit of Detection": The only requirement for this parameter is that it "should be reported". In our view, a comparison with min. or max. values given by the specifications would be more appropriate.

Finally, we believe that this document has been subjected to modifications that go far beyond simple editorial changes and it will certainly need an in-depth revision before it proceeds further in the Codex Step procedure.

Specific comments:

A new section "literature" should be added at the end of the document, indicating the references of the sources of the various formulas, examples and concepts.

It is not clear why the concept of "bias" and not that of "trueness" is used. Using "trueness" would be more consistent with the ISO-5725 series.

The use of measurement uncertainty to assess the fitness for purpose is not explained clearly enough.

For the estimate of the limit of quantification, the following formula is usually applied: LoQ = m+ 10 s.

If the formula proposed by the Working Group is maintained, the factor “k” in the formula should be clearly defined.

Finally, the examples given in Annex C are not very helpful. We would suggest that examples from the practice are taken and are considered in a step-by-step procedure and from A to Z using the proposed approach.