CODEX ALIMENTARIUS COMMISSION

E

Food and Agriculture Organization of the United Nations

World Health Organization

Viale delle Terme di Caracalla, 00153 Rome, Italy - Tel: (+39) 06 57051 - Fax: (+39) 06 5705 4593 - E-mail: codex@fao.org - www.codexalimentarius.org

Agenda Item 6 **CX/MAS 15/36/6**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

Thirty-sixth Session

Budapest, Hungary, 23 – 27 February 2015

DISCUSSION PAPER ON CRITERIA APPROACH FOR METHODS WHICH USE A 'SUM OF COMPONENTS'

(prepared by the United Kingdom[1])

BACKGROUND

1. At its 35th session of the Codex Committee on Methods of Analysis and Sampling (CCMAS) the Delegation of United States of America introduced the report of the electronic working group as presented in CX/MAS 14/35/5 and noted that there was general interest in the concept of developing criteria for Type I methods and/or multi-analyte methods, but that this was a starting point and no attempt was made to reach consensus on this. The Delegation highlighted the recommendations made and pointed out that the Committee would need to consider a number of factors when deciding on development of criteria for either Type I methods or for multi-analyte methods, such as: (i) when considering criteria for Type I methods, it may be possible to establish procedures for assessing equivalency between methods and not criteria. However, since not all Type I methods were created equal there may be instances where equivalency could not be established; (ii) in the case of multi-analyte methods, how to deal with TEFs, whether these should be left out of the standard as in the approach taken by CCFFP; and, (iii) whether a general approach was appropriate or whether different approaches would be necessary for multi-analyte methods (there might be differences between different toxins).

2. The Committee considered each of the recommendations.

Recommendation 1 – The establishment of Criteria for the different circumstances (Type I and multi-analyte method) should be addressed separately both during the development of the criteria and within the Procedural Manual

3. There was general agreement with this recommendation.

Recommendation 2 – Whether criteria for Type I methods should be established; or if a procedure for determining when methods have comparable performance should be developed; or if the current system should remain unchanged

4. There was general agreement that numerical criteria for Type I methods should not be developed, however procedures for establishing equivalency to Type I should be considered.

Recommendation 3 and 4 – establish a criteria approach for multi-analyte methods

5. There was general agreement that work should continue in this regard, that TEFs should not be contained within a specific analytical method and could be referenced either in the standard or elsewhere where they can be regularly updated and evaluated by internationally recognized procedures.

6. In view of the general discussion on the recommendations, the Committee agreed to pursue the work further through the establishment of two electronic working groups, open to all members and observers and working in English only, as follows:

[1] The lead of the EWG, Dr Andrew Damant

a. Development of procedures/guidelines for determining equivalency to Type I methods, led by United States of America, to prepare a discussion paper which would consider different approaches for different classes of Type I methods; and,

b. Development of a criteria approach for methods which use a "sum of components", led by United Kingdom. The working group would prepare a discussion paper that evaluates and discusses current options; and considers general guidelines and evaluates criteria for use on a case-by-case basis.

7. Also at the 35th Session of CCMAS the Committee discussed performance criteria for methods for the determination of marine biotoxins in the *Standard for Live and Raw Bivalve Molluscs.*

8. The Committee endorsed the criteria as proposed by CCFFP. The Committee noted that AOAC 2005.06 does not analyse all the substances in the table but covers major toxic components. It was also noted that it was helpful to provide to analysts information in the *Recommended Methods of Analysis and Sampling* (CODEX STAN 234-1999) on which methods of analysis meet the criteria.

9. The Committee endorsed AOAC 959.08 as well as AOAC 2011.27 (Receptor binding assay) as Type IV. The Committee was informed that AOAC 959.08 is not feasible in some countries where saxitoxin (STX) reference materials are not available, noting that its trade is restricted by the Chemical Weapons Convention.

10. At the 37th Session of the Codex Alimentarius Commission (CAC) the Commission further discussed Performance Criteria for Methods for the Determination of Marine Biotoxins (Section I-8.6) in the *Standard for Live and Raw Bivalve Molluscs*. The Commission considered the Draft section I-8.6 as endorsed and amended by the CCMAS.

11. However, there were concerns regarding the classification of the mouse bioassay (MBA) as Type IV which would mean that it could not be used for control, inspection and regulatory purposes in some countries. This would have a negative impact on trade as the method was widely used and efficient, and allowed for adequate protection of human health.

12. It was further noted that the criteria as described in the Procedural Manual were not applicable to biological methods, but rather to chemical methods and consideration should be given to exempt biological methods as currently was the case for PCR and ELISA methods.

13. Delegations reiterated their view that CCMAS should consider developing criteria for biological methods as the current criteria used for selection of methods applied to chemical methods, and led to the Type IV classification.

14. The Delegation of South Africa expressed a preference for adoption of both the biological and chemical methods rather than returning only the biological method to CCMAS.

15. It was noted that there was value in maintaining both the biological and chemical methods at the same status.

16. Other Delegations expressed the view that section I-8.6 allowed the use of both the MBA and chemical methods, and that CCMAS had followed the Principles for the Establishment of Codex Methods of Analysis. They also noted that CCMAS was in the process of addressing criteria for biological methods.

17. The CAC:

i. Adopted section I-8.6.1.

ii. Returned section I-8.6.2 to CCMAS with a request to review the typing of the methods in question and encouraged Members to submit information in order for CCMAS to take a decision on this matter.

iii. Encouraged CCMAS to proceed rapidly with its discussion on the way to deal with biological methods from a criteria approach perspective.

iv. Noted the reservation of South Africa to the decision in (ii) above.

18. As a result of discussions held at the 37th Session of the CAC the CCMAS working group tasked with developing a criteria approach for methods which use a "sum of components" was also asked to investigate the development of criteria for biological methods.

19. This discussion paper builds on work already co-ordinated and reported previously by the United States of America[2].

20. The eWG had over 60 participants. However, owing to the delay in preparing this discussion document the eWG have not yet been consulted so the paper presented is essentially a paper from the head of the UK delegation rather than the eWG *per-se*.

INTRODUCTION

21. The Procedural Manual establishes *General Criteria for the Selection of Methods of Analysis* (22nd Ed. 2014, English Version, p 68). Methods are evaluated on the characteristics of selectivity, accuracy, precision, limit of detection, sensitivity, practicability and applicability. It also allows for the establishment of other criteria as required and offers some guidance on choosing between different methods. The Procedural Manual also allows for the "Criteria Approach" as an alternative to the endorsement of a specific method (*ibid*). The Criteria Approach enables the establishment of a set of criteria (numeric values) which must be met by a method in order for the method to be applicable (i.e. "fit for purpose") to a specific standard. The Criteria Approach is applicable to fully validated Type II and III methods, except for methods such as PCR and ELISA, but it is not applicable to Type I methods. The Criteria Approach currently requires information on Applicability, Minimum Applicable Range, Limit of Detection and Quantitation, Precision (with criteria for reproducibility relative standard deviation), Recovery and Trueness (Procedural Manual 22nd Ed.2014, English Version pp 68). Two approaches for establishing criteria have been described in the Procedural Manual. The first utilizes the specified limit (maximum or minimum limit) to establish numeric criteria for the characteristics mentioned above and is summarized in Table 1. The second involves the conversion of a specific method to establish numeric criteria for the parameters listed in Table 1. Although the method should be validated and appropriate for the analyte and commodity, there is not a specific requirement that the method be endorsed prior to being "converted" to criteria. Although it is not specifically stated in the Procedural Manual, the *Guidelines for Establishing Numeric Values for Criteria* were developed considering only single analyte determinations. That is, methods where the concentration of a specific analyte is measured and that determination is assessed against a specification.

Table 1: Guidelines for establishing numeric values for the criteria.

Applicability:	The method has to be applicable for the specified provision, specified commodity and the specified level(s) (maximum and/or minimum) (ML). The minimum applicable range of the method depends on the specified level (ML) to be assessed, and can either be expressed in terms of the reproducibility standard deviation (s_R) or in terms of LOD and LOQ.
Minimum applicable range:	For ML ≥ 0.1 mg/kg, [ML - 3 s_R , ML + 3 s_R] For ML < 0.1 mg/kg, [ML - 2 s_R , ML + 2 s_R] s_R[3] = standard deviation of reproducibility
Limit of Detection (LOD):	For ML ≥ 0.1 mg/kg, LOD ≤ ML · 1/10 For ML < 0.1 mg/kg, LOD ≤ ML · 1/5
Limit of Quantification (LOQ):	For ML ≥ 0.1 mg/kg, LOQ ≤ ML · 1/5 For ML < 0.1 mg/kg, LOQ ≤ ML · 2/5

[2] CX/MAS 14/35/5 Discussion Paper on Considering Procedures for Establishing Criteria

[3] The s_R should be calculated from the Horwitz/Thompson equation. When the Horwitz/Thompson equation is not applicable (for an analytical purpose or according to a regulation) or when "converting" methods into criteria then it should be based on the s_R from an appropriate method performance study.

Precision: For ML ≥ 0.1 mg/kg, HorRat value ≤ 2
For ML < 0.1 mg/kg, the RSD_{TR} < 22% [44%?].
RSD_R[4] = relative standard deviation of reproducibility.
$RSD_R \leq 2 \cdot PRSD_R$

Recovery (R):

Concentration	Ratio	Unit	Recovery (%)
100	1	100% (100 g/100g)	98-102
≥10	10^{-1}	≥10% (10 g/100g)	98-102
≥1	10^{-2}	≥1% (1 g/100g)	97-103
≥0.1	10^{-3}	≥0.1% (1 mg/g)	95-103
0.01	10^{-4}	100 mg/kg	90-107
0.001	10^{-5}	10 mg/kg	80-110
0.0001	10^{-6}	1 mg/kg	80-110
0.00001	10^{-7}	100 µg/kg	80-110
0.000001	10^{-8}	10 µg/kg	60-115
0.0000001	10^{-9}	1 µg/kg	40-120

Trueness: Other guidelines are available for expected recovery ranges in specific areas of analysis.
In cases where recoveries have been shown to be a function of the matrix other specified requirements may be applied.
For the evaluation of trueness preferably certified reference material should be used.

22. The criteria in Table 1 must be approved for the determination in question.

SPECIFICATIONS REQUIRING A COMBINATION OF COMPONENTS

23. Although it is not specifically stated in the Procedural Manual, the *Guidelines for Establishing Numeric Values for Criteria* were developed considering only single analyte determinations. CCMAS paper CX/MAS 14/35/5 indicates the approaches detailed for single analytes in the Procedural Manual to be unsuitable for establishing criteria for specifications requiring the determination of a combination of components. For example, aflatoxins in nuts in the *General Standard for Contaminants and Toxins in Food and Feed* (CODEX STAN 193-2005) where the specification is for the concentration of total aflatoxin, which is determined as the sum of B1, B2, G1, and G2. Paper CX/MAS 14/35/5 extensively describes a number of possible options, each with benefits and drawbacks for establishing criteria in these situations. Namely,

- Option 2-1: Use the specification (sum of components) as the specified level (maximum/minimum limit) and develop numeric criteria based on this limit and the parameters listed in Table 1.
  - Option 2-2: Choose a suitable method and convert it into criteria using the guidelines currently listed in the Procedural Manual.
  - Option 2-2A: The numeric criteria are established from the approved method for each of the individual components.
  - Option 2-2B: The numeric criteria are established based on the specification and on the method performance for individual components.
- Option 2-3: Numeric criteria established based on the ML and the number of components.

24. A major problem central to all the options detailed within CX/MAS 14/35/5 is the determination of the predicted relative standard deviation ($PRSD_R$) criterion. The Horwitz/Thompson Equation was originally derived based on data associated with individual analytes and is not directly applicable to determining the $PRSD_R$ of a "sum of components." Therefore, the Horwitz/Thompson Equation or HorRat cannot be used to establish a numeric value for the precision. If one were to attempt to apply the Horwitz/Thompson Equation to the "sum of components" it could produce a situation where the precision of one or more individual component would need to exceed 100%. During discussions of

[4] The RSD_R should be calculated from the Horwitz/Thompson equation. When the Horwitz/Thompson equation is not applicable (for an analytical purpose or according to a regulation) or when "converting" methods into criteria then it should be based on the RSD_R from an appropriate method performance study.

CX/MAS 14/35/5 it was widely agreed that it was inappropriate to calculate the $PRSD_R$ value for summed component specifications from the ML value itself because in multi-component analysis the individual analyte measurements are correlated and therefore not independent. If the Horwitz/Thompson Equation is used then it should be restricted to individual analyte measurements.

25. In reality, for the majority of measurements undertaken for specifications which are summed components and the concentrations concerned fall into the Thompson $PRSD_R$ = 22% range so any method that does as well as 22% for the individual analyte should have acceptable precision. For higher levels the Horwitz value is likely to be the criterion.

26. A general question also raised within CX/MAS 14/35/5 was whether it is "permitted" within Codex to establish criteria for analytes that do not have associated specifications? Whilst this is a valid question this discussion paper takes the view that if individual analytes are specified (as is the case for aflatoxins in CODEX STAN 193-2005) then by default they are linked to the total specification and criteria can therefore be established.

27. Although not explicitly stated, paper CX/MAS 14/35/5 indicates the most pragmatic approach to be Option 2-3 where numeric criteria established are based on the ML and the number of components.

28. The Procedural Manual guidelines for establishing numeric values for LOQ are as follows:

Limit of Quantification (LOQ):	For ML ≥ 0.1 mg/kg, LOQ ≤ ML · 1/5
	For ML < 0.1 mg/kg, LOQ ≤ ML · 2/5

29. This is valid for analysing one component. When the ML is based on a sum of components, the LOQ for the individual component should theoretically be correspondingly low. When summing two components, the LOQ for each component should be the half for each component, and if summing three components; the LOQ for each component should be 1/3 of the LOQ.

30. Based on this, the following criteria for LOQ were suggested:

Limit of Quantification (LOQ):	For ML ≥ 0.1 mg/kg, LOQ ≤ ML · 1/5 · 1/*n*
	For ML < 0.1 mg/kg, LOQ ≤ ML · 2/5 · 1/*n*
	Where n = number of components

31. For multi-analyte analyses where all components are weighted equal, *n* is the number of components/analytes. The criteria for multi-analyte (and single analyte, *n*=1) would then be as given in Table 2.

Table 2: Guidelines for establishing numeric values for the criteria.

Applicability:	The method has to be applicable for the specified provision, specified commodity and the specified level(s) (maximum and/or minimum) (ML). The minimum applicable range of the method depends on the specified level (ML) to be assessed, and can either be expressed in terms of the reproducibility standard deviation (s_R) or in terms of LOD and LOQ.
Minimum Applicable	For ML/*n* ≥ 0.1 mg/kg, [ML/*n* - 3 s_R, ML + 3 s_R]
Range for <u>the individual components</u>[5]:	For ML/*n* < 0.1 mg/kg, [ML/*n* - 2 s_R, ML + 2 s_R]
	NB: the upper level is above the ML for the individual components.
Limit of Detection (LOD) for <u>the individual components</u>:	For ML/*n* ≥ 0.1 mg/kg, LOD ≤ ML/*n* · 1/10
	For ML/*n* < 0.1 mg/kg, LOD ≤ ML/*n* · 1/5
Limit of Quantification (LOQ) for <u>the individual components</u>:	For ML/*n* ≥ 0.1 mg/kg, LOQ ≤ ML/*n* · 1/5
	For ML/*n* < 0.1 mg/kg, LOQ ≤ ML/*n* · 2/5

[5] For multi-analyte analyses where all components are weighted equal, *n*=number of components/analytes.

Precision for the individual components: For ML/*n* ≥ 0.1 mg/kg, HorRat value ≤ 2

For ML/*n* < 0.1 mg/kg, the RSD_R < 44%.

RSD_R = relative standard deviation of reproducibility.

Recovery (R):	**Concentration**	**Ratio**	**Unit**	**Recovery (%)**
	100	1	100% (100 g/100g)	98-102
	≥10	10^{-1}	≥10% (10 g/100g)	98-102
	≥1	10^{-2}	≥1% (1 g/100g)	97-103
	≥0.1	10^{-3}	≥0.1% (1 mg/g)	95-103
	0.01	10^{-4}	100 mg/kg	90-107
	0.001	10^{-5}	10 mg/kg	80-110
	0.0001	10^{-6}	1 mg/kg	80-110
	0.00001	10^{-7}	100 µg/kg	80-110
	0.000001	10^{-8}	10 µg/kg	60-115
	0.0000001	10^{-9}	1 µg/kg	40-120

Trueness: Other guidelines are available for expected recovery ranges in specific areas of analysis. In cases where recoveries have been shown to be a function of the matrix other specified requirements may be applied. For the evaluation of trueness preferably certified reference material should be used.

Example A:

Aflatoxin, consisting of 4 analytes, B1, B2, G1 and G2, in peanuts.

The ML = 15 µg/kg,

As there are 4 analytes, n = 4,

ML/n = 15/4 µg/kg = 3.75 µg/kg

Using the excel spreadsheet on www.nmkl.org under "how to get method criteria based on ML", the following are established:

Minimum Applicable	0.002* - 0.022** mg/kg = 2 - 22 µg/kg
Range for the individual components:	*corresponding to ML = 3.75 µg/kg
	**corresponding to ML = 15 µg/kg
Limit of Detection (LOD) for the individual components:	0.75 µg/kg
Limit of Quantification (LOQ) for the individual components:	1.5 µg/kg
Precision for the individual components:	RSD_R ≤ 44%
Recovery (R):	40-120%

Examples on methods fulfilling the criteria:

AOAC 999.07 Immunoaffinity Column LX with post column derivatization

AOAC 2005.08 LC with Post-column photochemical derivatization

Examples on methods not fulfilling the criteria:

AOAC 975.36 (Romer minicolumn method) applicable for ≥ 10 µg/kg

AOAC 990.34 (Enzyme Linked Immunosorbent (ImmunoDot Screen Cup) Screening Assay ≥ 20 µg/kg

AOCS–AOAC 970.45, AOCS –AOAC 998.03. AOAC 993.17 Thin Layer Chromatography

Example B:

Biotoxins - Okadaic Acid (OA) Group (Assuming equally weighted components i.e. TEFs not applied).

The ML = 0.16 mg/kg[6]*,*

As there are 3 analytes, n = 3,

ML/n = 0.16/3 mg/kg = 0.05 mg/kg

Using the excel spreadsheet on www.nmkl.org under "how to get method criteria based on ML", the following are established:

Minimum Applicable	0.03* - 0.26** mg/kg
Range for the individual components:	*corresponding to ML = 0.05 mg/kg
	**corresponding to ML = 0.16 mg/kg
Limit of Detection (LOD) for the individual components:	0.01 mg/kg
Limit of Quantification (LOQ) for the individual components:	0.02 mg/kg
Precision for the individual components:	$RSD_R \leq 44\%$
Recovery (R):	60-115%

Examples on methods fulfilling the criteria:

None

Examples on methods not fulfilling the criteria:

European Union Reference Laboratory Method for Marine Biotoxins, OA and AZA SOP, 2011 – the LOQ of the method is interpreted as being 0.04 mg/kg for both OA and AZA.

Example C:

Biotoxins - Saxitoxin (STX) Group (Assuming equally weighted components i.e. TEFs not applied).

The ML = 0.8 mg/kg[7]*,*

As there are 15 analytes, n = 15,

ML/n = 0.8/15 mg/kg = 0.05 mg/kg

Using the excel spreadsheet on www.nmkl.org under "how to get method criteria based on ML", the following are established:

Minimum Applicable	0.03* - 1.2** mg/kg
Range for the individual components:	*corresponding to ML = 0.05 mg/kg
	**corresponding to ML = 0.8 mg/kg
Limit of Detection (LOD) for the individual components:	0.01 mg/kg
Limit of Quantification (LOQ) for the individual components:	0.02 mg/kg
Precision for the individual components:	$RSD_R \leq 44\%$
Recovery (R):	60-115%

[6] Officially the ML is 0.16 mg/kg of okadaic equivalent but for the purposes of this example the TEFs have not been taken into account.

[7] Officially the ML is 0.8 mg/kg (2HCL) of saxitoxin equivalent but for the purposes of this example the TEFs have not been taken into account.

Examples on methods fulfilling the criteria:

AOAC 2005.06 Paralytic Shellfish Poisoning Toxins in Shellfish – the LOQ of the method is interpreted as being 0.02 mg/kg for STX and 0.008 mg/kg for dcSTX.

Examples on methods not fulfilling the criteria:

AOAC 2005.06 Paralytic Shellfish Poisoning Toxins in Shellfish – the LOQ of the method is interpreted as being 0.125 mg/kg for both GTX 2,3 (together) and 0.03 mg/kg for B-1.

23. Whilst the methods detailed in Examples B and C each have acceptable precision the methods do not fulfil criteria in terms of LoQ for each analyte. This demonstrates a principle problem when the LoQ (and LoD) is prorated against the number of analytes (*n*). The greater the number of analytes the lower the LoQ required if the specification relates to a sum of components.

24. Option 2-2A of CX/MAS 14/35/5 describes how numeric criteria may be established from the approved method for each of the individual components. Annex 2 of CX/MAS 14/35/5 states that the RSD_{total} becomes smaller when the number of components increases. However, Annex 2 did not take into account correlation and covariance effects where the following example D[8] illustrates the importance of such issues.

Example D,

Table 3 shows results from the analysis of aflatoxins B1, B2, G1 and G2 in foodstuffs where several laboratories have reported separate results for the four aflatoxins in a particular material. Their variances and standard deviations are also shown. Can we estimate the standard deviation for total aflatoxin directly from these four standard deviations?

Table 3: Replicated results from the determination of aflatoxins (ppb mass fraction).

Laboratory	B1	B2	G1	G2	Total
1	8.5	4.3	3.5	1.6	17.9
2	4	2.5	1.7	2.1	10.3
3	6.6	3.6	2.1	2	14.3
4	5.9	3.4	2.3	2.2	13.8
5	4.2	2.2	1.8	1.6	9.8
6	6.2	3.5	2.6	2.7	15.0
7	7.1	3.8	2.6	2.5	16.0
8	5.2	3.4	2.1	2.2	12.9
9	4.9	2.45	2.15	1.8	11.3
10	6.3	3.3	2.3	1.9	13.8
Variance	1.881	0.438	0.259	0.129	6.40
Standard deviation	1.371	0.662	0.509	0.359	2.53

If we assumed that the results for the four aflatoxins are uncorrelated we would take the standard deviation of the sum of the four results to be the square root of the sum of the individual variances, namely:

$$\sqrt{1.881 + 0.438 + 0.259 + 0.129} = 1.65$$

[8] RSC AMC Technical Brief No. 30, 2008. The standard deviation of the sum of several variables (http://www.rsc.org/Membership/Networking/InterestGroups/Analytical/AMC/TechnicalBriefs.asp)

But if we calculate the individual total aflatoxin results for the laboratories, we get values shown in the last columns of Table 3. These have a standard deviation of 2.53, considerably larger than the calculation above so why is there a difference?
The values differ because the observations are not independent; they show appreciable correlation. This can be seen by calculating their covariances cov(x,y) (Table 4) and the related correlation coefficients r(x,y) (Table 5), using the formulae:

$$cov(x,y) = \sum_i (x_i - \hat{x})(y_i - \hat{y})/(n-1)$$
$$\equiv r(x,y)s_x s_y$$

Where x_i, y_i are the i-th pair of variables x, y and s_x, s_y are individual standard deviations.

Table 4: Covariance matrix

	B1	B2	G1	G2
B1	**1.881**	0.848	0.635	0.032
B2	0.848	**0.438**	0.277	0.066
G1	0.635	0.277	**0.259**	0.000
G2	0.032	0.066	0.000	**0.129**

Table 5: Correlation Coefficients.

	B1	B2	G1	G2
B1	1	0.934	0.909	0.064
B2	0.934	1	0.823	0.276
G1	0.909	0.823	1	0.000
G2	0.064	0.276	0.000	1

Several of the correlation coefficients are well over 0.5 indicating that they will substantively affect the combined standard deviation.

With some correlation between the variables, as we have in our example data, the correct standard deviation of the sum is the square root of the sum of the variances and the covariances, that is:

$$\sqrt{(1.881 + 0.438 + 0.259 + 0.129 + 2(0.848 + 0.635 + 0.032 + 0.277 + 0.066 + 000))} = 2.53$$

As we saw above, this result can be obtained directly as the simple standard deviation of the calculated total aflatoxin contents above. (Notice that both cov(x,y) and cov(y,x) have to be included in the sum, hence the factor of 2 for the covariance terms).

Exactly the same principles as above apply when estimating the standard uncertainty for sums of variables. The standard deviation above provides the standard uncertainty associated with random effects for the total aflatoxin content reported by a single laboratory. For an average of n results (for example, from a series of observations within a single-laboratory run), the calculated standard deviation should be divided by $\sqrt{n}$.

If there is any correlation between variables which are to be added together (or, indeed, combined in any way), it is important to take proper account of that correlation in estimating the standard deviation of the result. If the raw data are available, this can either be done by calculating the individual results and taking their standard deviation or by calculating the necessary covariances and summing those. If only the standard deviations and correlation coefficients (or covariances) are available, it is necessary to calculate the combined standard deviation from the covariances.

25. Example D shows one approach to determine the measurement (standard) uncertainty of an analytical result based upon the sums of components but there are others published in the literature[9,10].

26. Section 2 (page 85) of the *Procedural Manual* states, *"An allowance is to be made for the measurement uncertainty when deciding whether or not an analytical result falls within the specification. This requirement may not apply in situations when a direct health hazard is concerned, such as for food pathogens.* The *Guideline on Measurement Uncertainty* (CAC/GL 54-2004) provides general guidance on measurement uncertainty but does not cover the issue of measurement uncertainty associated with analytical values that are themselves sums of components. The *Guidelines on Estimation of Uncertainty of Results* (CAC/GL 59-2006) provides guidance on the estimation of uncertainty of results within the area of pesticide analysis and states, "The estimation of uncertainty of results for multi-component residues arising from the application of technical mixtures including structural and optical isomers, metabolites and other breakdown products may require a different approach particularly where the MRL has been established for the sum of all or some of the component residues. The assessment of the random and systematic errors of the results based on the measurements of multiple peaks is explained in detail in a recent publication[11].

27. Whilst not explicitly linked to the rationale behind this paper the issues raised in Examples B, C and D, and the preceding paragraphs, indicate a need for the development of guidance to explain approaches that may be taken when determining the measurement uncertainty associated with a reported analytical result that is based upon the summation of components.

28. Whilst reservations were expressed in CX/MAS 14/35/5 about the approach given within Option 2-2A the approach was taken in the *Standard for Live and Raw Bivalve Molluscs (*CODEX STAN 292-2008) (Revised 2014) where numeric criteria (based upon advice from CCMAS) were established from the approved methods for each of the individual components in each of the toxin group. As such precedence has now been set. The number of analytical areas within the Codex framework where a sum of components approach needs to be taken is limited. Consequently, and also owing to difficulties expressed in previous paragraphs, it is recommended that if numerical criteria for such methods needs to be established then they are developed from the approved methods for each individual component and not theoretically from the ML value.

TOXIC EQUIVALENCE FACTORS

29. For certain commodities or analytes there are specifications where the individual concentrations of multiple analytes are determined by a single method, the concentrations are converted to a "toxic equivalent" using a toxic equivalency factor (TEF) and the specification is a limit based on the sum of equivalents. One example of this approach is the determination of the Saxitoxin group in the *Standard for Live and Raw Bivalve Molluscs* (CODEX STAN 292-2008). The specification is for the concentration of saxitoxin equivalents which is determined from 12 saxitoxin congeners[12] each multiplied by a TEF and summed. TEFs are also used in other determinations, such as dioxins and dioxin-like PCBs, and PAHs. The current Criteria Approach in the Procedural Manual was not developed considering specifications which use TEF or a sum of toxic equivalents.

30. The use of a TEF to determine a "toxic equivalent" requires a calculation, and if this calculation is part of the method, then historically CCMAS would consider such methods as Type I. Even if the analytical procedure to determine the value prior to conversion was rational (Type II/III), the final

---

[9] Gauthier Eppe, Gianfranco Diletti, Alwyn Fernandes, Johannes Haedrich, Jerry Hart, Helge Hove, Anna Laura Iamiceli, Alexander Kotz, Rainer Malisch, Philippe Marchand, Wolfgang Moche, Georges Scholl, Giampiero Scortichini, Thorsten Bernsmann, Yves Tondeur and Wim Traag. Measurement Uncertainty for Persistent Organic Pollutants by Isotope-Dilution Mass Spectrometry. Paper presented at Dioxins 2014 (In press?).

[10] Medina-Pastor, P., Valverde, A., Pihlstrom, T., Masselte, S., Gamon, M., Mezcua, M., Rodriguez-Torreblanca, C. and Fernandez-Alba, A.R. Comparative study of the main top-down approaches for the estimation of measurement uncertainty in multiresidue analysis of pesticides in fruits and vegetables. *J. Agric. Food Chem.* **59**, 7609-19, 2011.

[11] Soboleva E., Ambrus A., Jarju O., Estimation of uncertainty of analytical results based on multiple peaks, *J. Chromatogr. A*. **1029**, 161-166, 2004.

[12] There are more than 12 saxitoxin congeners identified, however the currently endorsed method (AOAC 2005.05) only lists 12 compounds.

determination is Type I because the calculation is empirical. A possible alternative to including the TEFs in the method would be to include them in the standard.

31. Where TEFs are set these normally relate to the most toxic substance within the suite of substances being analysed so a pragmatic approach would be to ensure that any analyte specific method performance criteria stipulated are based upon the most toxic component itself (e.g. for PSP shellfish toxins, STX, Oshima) as this will be at the lowest mass fraction and therefore have the poorest expected precision. If Horwitz's model applies ideally, all the others would be more precise if they were the dominant component in a test sample at the limit because their mass fractions would be higher. It follows that if a method works when the most toxic component is at the toxic equivalent limit, it should work better when any other weighted sum is at the limit.

32. During discussions of CX/MAS 14/35/5 at CCMAS35 it was widely agreed that TEFs should not be contained within a specific analytical method, but should be captured in the Standard. This was the approach taken when amended *Standard for Live and Raw Bivalve Molluscs* (CODEX STAN 292-2008) where un-weighted numerical performance criteria (i.e. TEFs not applied) were established from the various approved methods.

NUMERICAL METHOD PERFORMANCE CRITERIA FOR BIOLOGICAL METHODS

33. CCMAS has historically taken the view that biological methods such as the mouse bioassay should be classified as being Type I (Defining Method) where the method determines a value that can only be arrived at in terms of the method *per-se* and serves by definition as the only method for establishing the accepted value of the item measured).

34. Table 6 shows the results of a keyword search of the *Recommended Methods of Analysis and Sampling* (CODEX STAN 234-1999) for 'bioassay' methods and the awarded method type.

Table 6: 'Bioassay' Methods Detailed Within CODEX STAN 234-1999.

Commodity	Provision	Method	Principle	Type
Margarine	Vitamin D	AOAC 936.14	Bio-assay (Rat)	II
Minarine	Vitamin D	AOAC 936.14	Bio-assay (Rat)	II
Special foods	Folic acid	AOAC 944.12	Micro-bioassay	II
Special foods	Nicotinamide for milk-based foods	AOAC 944.13	Micro-bioassay	II
Special foods	Pantothenic acid/enriched foods	AOAC 945.74	Micro-bioassay	II
Special foods	Pantothenic acid/non-enriched foods	The Analyst 89 (1964):1, 3-6, ibid. 232 US Dept Agr., Agr. Handbook 97 (1965)	Micro-bioassay	IV
Special foods	Protein efficiency ratio (PER)	AOAC 960.48	Rat bio-assay	I
Special foods	Vitamin B12	AOAC 952.20	Micro-bioassay	II
Special foods	Vitamin B6	AOAC 961.15	Micro-bioassay	II
Special foods	Vitamin D	AOAC 936.14	Rat bio-assay	IV
Follow-up formula	Pantothenic acid	AOAC 992.07 (Measures total pantothenate (free pantothenic acid + CoA- + ACP-bound) and measured as D-pantothenic acid (or calcium D-pantothenate)	Micro-bioassay	II
Infant formula	Folic acid	AOAC 992.05 (Measures free folic acid + free, unbound natural folates, aggregated and measured as folic acid) EN 14131:2003 (Total folate (free + bound),	Micro-bioassay	II

		aggregated and measured as folic acid)		
Infant formula	Niacin	AOAC 985.34 (niacin (preformed) and nicotinamide)	Micro-bioassay and turbidimetry	III
Infant formula	Vitamin B6	AOAC 985.32	Micro-bioassay	III
Infant formula	Vitamin B6	EN 14166:2008 (Aggregates free and bound pyridoxal, pyridoxine and pyridoxamine and measures as pyridoxine)	Micro-bioassay	III

35. Although method AOAC 936.14 is labelled as 'bioassay' within CODEX STAN 234 it is in fact a rat bio-assay method. The rat bio-assay methods detailed in CODEX STAN 234 are classified and being Type I, Type II (Reference Method) or Type IV (Tentative Method) methods. At the 37th Session of CCMAS the mouse bioassay (AOAC 959.08) and the receptor binding assay (AOAC 2011.27) were both classified as being Type IV which generated subsequent discussion at the 37th Session of the CAC.

36. From the above information it is unclear whether a consistent approach is being taken to the 'typing' biological methods such as the mouse or rat bio-assay. At the 36th Session of CCMAS it is recommended that CCMAS consider to formation of an eWG to investigate the 'typing' of mouse and rat bio-assays in more detail with the aim of developing a harmonised and consistent approach.

37. If it is considered that all mouse and rat bio-assay methods are Type I then the issue of numerical performance criteria is irrelevant because the methods are Defining Methods and therefore have no equivalents.

38. If it is considered that mouse and rat bio-assay methods can be classified as being Type II then the approach detailed in Option 2-2A of CX/MAS 14/35/5 (i.e. numerical performance criteria are established from the approved method) should be adopted in order to maintain consistency with the approach taken in Codex Standard 292-2008. If mouse and rat bio-assay methods can be classified as being Type II then work needs to be potentially undertaken by an eWG to investigate equivalence between biological and chemical methods.

TENTATIVE RECOMMENDATIONS AND POINTS FOR DISCUSSION

1) If numeric performance criteria need to be established for methods applied to provisions that involve the summation of analytical components then these should be established from the approved method for each of the individual components rather than the ML.
2) If numeric performance criteria are established for provisions that require the summation of components then they should be established for each of the components on an un-weighted basis.
3) If recommendations 1-2 are accepted then the *General Criteria for the Selection of Methods of Analysis* section of the Procedural Manual should be amended to indicate that the process based upon the ML value is only suitable for single-analyte analyses and that if criteria need to be developed for provisions that require a sum of components then the process based upon the approved method is the preferred option.
4) If provisions require the use of toxic equivalence factors (TEFs) then these should be detailed within the Codex specification but should be independent from the development of numeric performance criteria.
5) The remit of the 'sum of components' eWG should be expanded to include the development of a discussion paper on approaches which can be used to establish the measurement uncertainty of analytical results that are sums of components.
6) CCMAS should consider the formation of an eWG to investigate on-going issues with the 'typing of mouse and rat bio-assay methods and their equivalence to chemical based analyses.