Available online at www.sciencedirect.com
SCIENCE DIRECT®
Analytica Chimica Acta 519 (2004) 243–253
ANALYTICA CHIMICA ACTA
www.elsevier.com/locate/aca

Performances and limitations of the HRMS method for dioxins, furans and dioxin-like PCBs analysis in animal feedingstuffs Part II: Does it comply with the European pro-active approach?

Gauthier Eppe[a,*], Guy Maghuin-Rogister[b], Edwin De Pauw[a]

[a] *Mass Spectrometry Laboratory, C.A.R.T., Department of Physico-Chemistry, University of Liege, Allee de la chimie B6c Sart-Tilman, B-4000 Liege, Belgium*
[b] *Laboratory of Analysis of Foodstuffs of Animal Origin, C.A.R.T., University of Liege, B-43bis Boulevard de Colonster 20, B-4000 Liege, Belgium*

Received 7 March 2004; received in revised form 9 June 2004; accepted 9 June 2004

Available online 17 July 2004

Abstract

Based on the results obtained from the inter-laboratory study in Part 1, different approaches for detection and quantification limits are evaluated and discussed. An overview of the most commonly used concepts and terminologies in analytical chemistry is presented with the aim of establishing a link between them. Whatever the method used by laboratories for detection limit assessment, the median LOD value reported for the less chlorinated PCDD/Fs (i.e. 0.02 ng/kg) is in good agreement with the values recalculated using the inter-laboratory data. For LOQ, the Eurachem approach based on a pre-established percentage of repeatability RSD appears to be suitable. The study shows that a pre-established RSD_r of 20% is recommended in order to achieve an acceptable LOQ of 0.05 ng/kg per congener. The 20% value seems to be sufficiently low to get tolerable RSD close to maximum limits. Furthermore, the repeatability and the reproducibility standard deviation against parts-per-trillion congener levels has been modeled by inverse first order functions. This congener precision model provides an interesting tool to subsequently assess the performances of the method in TEQ close to regulatory limits. Finally, the paper discusses two different ways of reporting and interpreting the results to assess compliance against statutory limits.

Keywords: Dioxins; Furans; Dioxin-like PCBs; HRMS; Detection and quantification limits; Compliance and non-compliance; Measurement uncertainty

1. Introduction

After several dioxin crises in the food chain in Europe, the need to set regulatory limits in various food and feed matrices is easily understandable. For the study concerned, the maximum level was set at 0.75 ng WHO-PCDD/F TEQ/kg of product for feed materials of plant origin, while the action level was established by reducing the maximum levels by at least 30% (i.e. 0.5 ng WHO-PCDD/F TEQ/kg of product for feed materials of plant origin) with the aim of triggering actions on identification of sources and pathways of dioxin contamination. The ultimate goal is to establish target levels to reduce the exposure of the majority of the European population below the tolerable weekly intake (TWI). As can be seen, the legislation established regulatory levels in toxic equivalence (TEQ) units in order to facilitate risk assessment. Based on toxicological considerations, the trend is to decrease the contamination levels, whereas the number of chemical compounds concerned in the so-called TEQ approach will increase in the near future, due to the addition of dioxin-like (D-L) PCBs. By decreasing statutory levels to such low levels, does it mean that the analytical consequences can lead to a situation in which the best available technology to provide reliable results (as the state-of-the-art HRMS method) will be the only acceptable technology? In that case, the European legislation would be inconsistent with the monitoring approach, which promotes the use of screening methods. Part 2 will concentrate on the analytical performances that laboratories are able to realize with the current HRMS technique and will present the issue from a

* Corresponding author. Tel.: +32-4-366-3422; fax: +32-4-366-4387.
E-mail address: g.eppe@ulg.ac.be (G. Eppe).

0003-2670/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.aca.2004.06.021

different angle, i.e. the analytical point of view for the future establishment of target levels.

Part 1 of this paper focused on the performances of the HRMS method for dioxin, furans and dioxin-like PCBs in animal feedingstuffs. Inter-laboratory results allowed evaluating different performance parameters of the method, such as accuracy, by assessing precision and trueness. Precision was evaluated afterwards by characterizing the two extreme measures of precision: repeatability and reproducibility.

For ultra-trace analysis, method validation also requires evaluating many other fundamental performance characteristics. Among these, detection limits and quantification limits are certainly considered as key issues for the determination of PCDD/Fs at sub-parts per trillion levels in food and feed, particularly since the European legislation introduced the concept of lower bound, middle bound and upper bound values for reporting [1]. Unfortunately, different concepts and terminologies for detection and quantification limits abound in the chemical literature, underlining decades of confusion [2]. The need for a unified approach for detection and quantification limits was highlighted and reflected by the publications of the IUPAC recommendations [3] and later by the ISO 11843 [4]. The UNEP POP workshop [5] has also recently emphasized the need for greater consensus on the definitions for detection limits and reporting methods. The common approach is based on hypothesis tests characterized by a rate of false positive (α, Type-1 error) and a rate of false negative (β, Type-2 error), set according to the confidence level. On the other hand, the 'dioxin scientific community' generally uses the terminology, limit of detection (LOD) and limit of quantification (LOQ), and mainly applies the 'signal/noise' approach as described in the Standard EN1948-3 [6], whereas the end-users of the EPA method 1613 revision B [7] have to determine a method detection limit (MDL) and a maximum level (ML).

More recently, the European Commission laid down a decision of the Commission 2002/657/EC [8] regarding the performances of analytical methods and the interpretation of the results for residue analysis. The concept of the decision limit (CC_α) defined as 'the limit at and above which it can be concluded with an error of probability of α that a sample is non-compliant' and the detection capability (CC_β) defined as 'the smallest content of the substance that may be detected, identified and/or quantified in a sample with an error of probability β' were introduced. In fact, these concepts result from the IUPAC recommendations and the ISO 11843 mentioned above. The 'negative and positive' terminology has been abandoned and replaced respectively by the terms, 'compliant or non-compliant'. Another interesting feature of the CC_α and the CC_β concept is that for substances for which permitted limits (PLs) have been established, as dioxin compounds, they can be used to judge for compliance or non-compliance. CC_α and/or CC_β give values above the PL for which, at that level, a decision of non-compliance becomes extremely likely. Currently, there are mainly two schools in Europe regarding the reporting and interpretation of results for residues in food and feed control purposes. One is to use CC_α and CC_β values above the PL for deciding whether an analytical result falls within the specification or not; the other is to estimate the measurement uncertainty associated with an analytical result close to the maximum level. The result should take the form of 'a ± U' where 'a' is the best estimate of the true value of the concentration of the measurand and 'U', the expanded uncertainty. The inter-laboratory study carried out according to ISO 5725 in Part 1 was a precious source of data to support uncertainty estimations. The remaining significant sources of uncertainty, not covered by the collaborative study, also have to be evaluated. Another important issue is also raised when uncertainty estimation has to be reported at the background level, as is generally the case for dioxin and furan congeners levels in animal feedingstuff samples. The paper also discusses how to deal with the evaluation of uncertainty for non-detected congeners and how to report them.

Finally, an overview of the different approaches introduced here-above for dioxin and furan congeners in animal feedingstuffs is evaluated and discussed.

2. Detection and quantification limits for dioxin and furan congeners

2.1. Definitions

In broad terms, the limit of detection is the smallest concentration level that can be determined statistically different from a blank at a specified level of confidence [9]. This corresponds to the critical value L_C defined by Currie [3] or X_C (critical value of the net state variable) defined in ISO 11843 [4], or more recently for CC_α [8]. These terms are commonly used in analytical chemistry. The approach is based on hypothesis tests where two risks values α (type-1 error) and β (type-2 error) are defined as follows: α corresponds to the risk of detecting an analyte, although it is not present, also called a false positive or a false non-compliant decision. β corresponds to the risk of not detecting an analyte while it is present, i.e. a false negative or a false compliant decision. At the L_C level, only the decision of 'detected' or 'non detected' is made. The rate of β error is therefore 50%. Instead of LOD, EPA preferred to use the term 'method detection limit' (MDL), which is the minimum concentration of a substance that can be measured and reported at a level of confidence that the analyte concentration is greater than zero. It is determined from analysis of a sample in a given matrix containing the analyte. The MDL takes into account the whole analytical process. It is important to distinguish it from the instrument detection limit (IDL), which is the smallest signal above background that an instrument can detect. IDL involves only one component of the analytical process.

Keith [9] also introduced the reliable detection limit (RDL), which is the concentration level at which a decision

is extremely likely. This corresponds to the detection limit L_D defined by Currie [3] or X_D (the minimum detectable value of the net state variable) [4] or more recently, for CC_β [8]. L_D, X_D, CC_β or RDL are determined at levels beyond L_C, X_C or CC_α where the risk of error β becomes acceptably low. For dioxin analysis, it is recommended to set α and β at 5% (i.e. substances B in annex 1 of directive 96/23/EC) [10]. Assuming a normal distribution of the data (an assumption which is at least questionable at this level of concentration) and a constant standard deviation in the range between L_C and L_D, then $L_C = 1.645\sigma_B$ (where $\alpha = 5\%$ and $\beta = 50\%$) and $L_D = 3.29\sigma_B$ (where $\alpha = 0.05\%$ and $\beta = 5\%$), σ_B being the standard deviation of the blank. After rounding-off, this led to the famous definition of $L_D \approx 3\sigma_B$.

The limit of quantification (LOQ), also frequently reported in the literature as 'limit of determination', is the smallest quantity or concentration that can be quantified with a given level of confidence. LOQ, L_Q or X_Q is different, and more difficult, than measuring the presence or absence of an analyte. It is an indicative value and is sometimes defined at a precision level (as a percentage of RSD, commonly 10%) arbitrarily fixed (Eurachem approach) [11]. Keith [9] recommended an LOQ $\approx 10\sigma_B$, corresponding to an uncertainty of ±30% at a 99% confidence level. Frequently, LOQ is also simply defined as a multiple of the LOD (usually by a factor between 2 and 3). The signal to noise (*S*/*N*) ratios are still a widely used method for chromatographic techniques.

For dioxin and furan congeners, the term 'maximum limit of quantification' (ML) has been proposed in the EPA 1613 method revision B [7]. The ML equates to a limit of quantification and not a limit of detection (or MDL), as it is generally used in EPA documents. The ML for each PCDD/F congener is defined as the level at which the entire analytical system must give a recognizable signal and acceptable calibration point. It is equivalent to the concentration of the lowest calibration standard, assuming that all method-specified sample weights, volumes and clean-up procedures have been employed. Laboratories have to demonstrate that the MDL is lower than one-third of the ML. The ML is calculated based on inter-laboratory analyses of the analyte in the matrix of concern, meeting pre-established acceptance criteria of precision and accuracy. In Europe, the Standard EN 1948-3 [6] emphasized the use of *S*/*N* ratio approach for LOQ as well as the recent recommendations from the expert committee for Commission Directives 2002/69/EC and 2002/70/EC. The LOQ is defined here as the concentration of an analyte which produces an instrumental response at two different ions to be monitored with a *S*/*N* ratio of 3:1 for the less sensitive signal and fulfillment basics requirements such as, for example, retention time, isotope ratio. Obviously, the noise should be the noise from the whole analytical process. Otherwise it is called an IDL. Furthermore, end-users know that with 'real noise' it is sometimes difficult to simultaneously fulfill the retention time and the isotopic ratio requirements at *S*/*N* of 3:1. Moreover, how should one deal with LOD/LOQ for congeners that are systematically present in blank procedure [12] (like hepta-octa CDD/F congeners or some D-L PCBs)? This definition might then lead, in some cases, to 'an optimistic value' of quantification limits.

For the sake of clarity, in the rest of the paper, the terms LOD/LOQ (which are the terms commonly used by the 'dioxin scientific community') correspond to L_D, X_D, CC_β, RDL and L_Q, X_Q, ML, respectively.

2.2. LOD/LOQ reported by laboratories

As already mentioned in Part 1, participants were asked to report their detection limit and quantification limit per congener. No information was collected on the way they calculated or evaluated them. It was decided to calculate the median values to represent one detection limit and one quantification limit per congener. As scarce data were provided for D-L PCBs, this paper focuses only on dioxin and furan congeners. Table 1 summarizes the median LOD/LOQ of the HRMS method for the 17 congeners. The range of LOD/LOQ reported by the laboratories is indicated in brackets. In general, for the less chlorinated congeners (tetra, penta and hexa-CDD/Fs), laboratories reported similar LOD/LOQ per congener. It has therefore been assumed that one LOD and one LOQ representing those congeners can be assessed (i.e. 0.02 ng/kg for LOD and 0.05 ng/kg for LOQ). On the other hand, the LOD/LOQ for the hepta-octa-CDD/F congeners were certainly penalized by their presence in blank sample, and consequently higher values were reported. The sum, expressed in TEQ, gives respectively 0.07 ng WHO-TEQ/kg for LOD and 0.17 ng WHO-TEQ/kg for LOQ. Directive 2002/70/EC specifies that the HRMS method should have a sum of LOQ in TEQ calculation in the range of about one-fifth of the maximum level (i.e. 0.75 ng WHO-TEQ/kg) in order to achieve an

Table 1
Median LOD and median LOQ of the HRMS method for PCDD/F analysis in feed

Compounds	Median LOD (ng/kg)	Median LOQ (ng/kg)
2,3,7,8-TCDD	0.020 [0.005–0.04]	0.050 [0.01–0.080]
1,2,3,7,8-PeCDD	0.020 [0.005–0.04]	0.050 [0.01–0.080]
1,2,3,4,7,8-HxCDD	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,6,7,8-HxCDD	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,7,8,9-HxCDD	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,4,6,7,8-HpCDD	0.180 [0.013–0.35]	0.360 [0.025–0.63]
1,2,3,4,6,7,8,9-OCDD	0.160 [0.025–0.40]	0.320 [0.050–2.56]
2,3,7,8-TCDF	0.020 [0.005–0.07]	0.050 [0.010–0.14]
1,2,3,7,8-PeCDF	0.020 [0.005–0.04]	0.050 [0.010–0.08]
2,3,4,7,8-PeCDF	0.020 [0.005–0.04]	0.050 [0.010–0.08]
1,2,3,4,7,8-HxCDF	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,6,7,8-HxCDF	0.020 [0.008–0.12]	0.050 [0.015–0.24]
2,3,4,6,7,8-HxCDF	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,7,8,9-HxCDF	0.020 [0.008–0.12]	0.050 [0.015–0.24]
1,2,3,4,6,7,8-HpCDF	0.030 [0.013–0.25]	0.060 [0.025–0.46]
1,2,3,4,7,8,9-HpCDF	0.040 [0.013–0.18]	0.080 [0.025–0.36]
1,2,3,4,6,7,8,9-OCDF	0.090 [0.025–0.24]	0.180 [0.050–0.42]
Total-TEQ PCDD/Fs	0.070	0.170

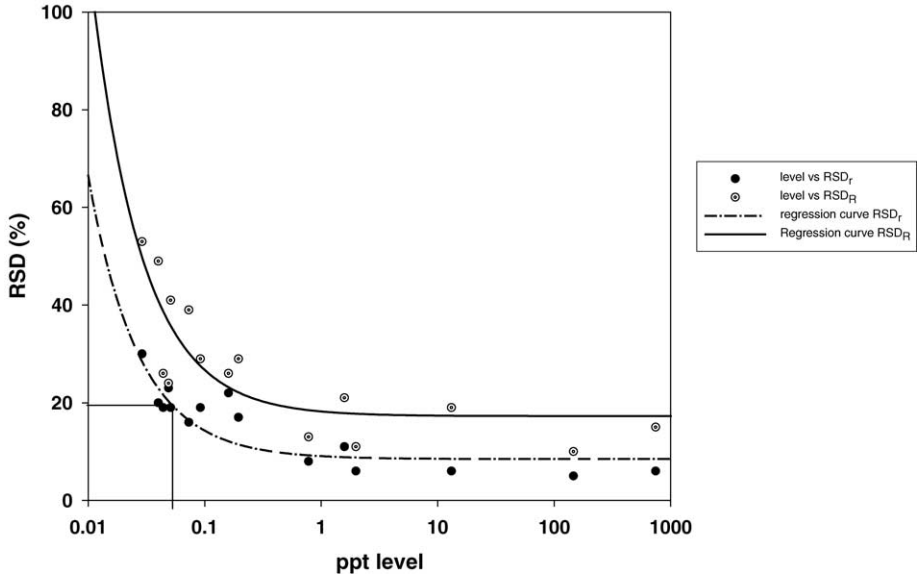

Fig. 1. RSDs vs. mean level (ppt) for PCDD/F congeners.

acceptable CV at the level of interest. The values reported in Table 1 seem to be sufficiently low to comply with the criteria. However, the question that must now be asked is the following: are those values consistent and realistic?

2.3. The IUPAC approach

One of the main conclusions drawn from Part 1 was, that in this particular study, the performances of the HRMS method seem to be congener-independent for repeatability and reproducibility over a concentration range spanning more than four orders of magnitude. Fig. 1 shows the precision of the method for PCDD/F congeners, except for 2,3,7,8 TCDD, for which the ISO 5725 procedure was not effective and 1,2,3,7,8,9 HxCDF, for which too scarce results were reported. The precision obtained using ISO 5725 Standard was modeled by the following equations:

$$\mathrm{RSD}_{\mathrm{r}i}(\%) = 0.58c_i^{-1} + 8.2 \quad (1)$$

$$\mathrm{RSD}_{\mathrm{R}i}(\%) = 1.69c_i^{-1} + 11.9 \quad (2)$$

Where c_i is the level of the congener i stated in parts-per-trillion.

The test material used for the collaborative study was naturally contaminated with PCDD/Fs and among them, six congeners had consensus mean values below 0.08 ppt (see Fig. 1), i.e. in the range of LOD/LOQ reported by the participants. Information provided from the inter-laboratory data in this range are extremely relevant. In this delimited region close to detection limit, the repeatability standard deviation (S_r) increases in a linear fashion with sub-ppt level, as can be seen in Fig. 2. The linear dependency of S_r with level is called 'heteroscedasticity' and represented here by the following equation:

$$S_\mathrm{r} = 0.0055 + 0.0868\,c \quad (\mathrm{ng/kg}) \quad (3)$$

The y-intercept represents the estimation of the standard deviation of the blank (i.e. $S_\mathrm{B} = 0.0055$ ng/kg). The corresponding critical level L_C is therefore equal to $1.645S_\mathrm{B}$, which gives 0.009 ng/kg per congener at the 95% confidence level. The detection limit $L_\mathrm{D} = 1.645S_\mathrm{B} + 1.645S_\mathrm{D}$ (where $S_\mathrm{D} = 0.0055 + 0.0868L_\mathrm{D}$). Thus L_D equals to $3.84S_\mathrm{B}$, which gives 0.021 ng/kg per congener at a 95% confidence level. The detection limit of 0.021 ng/kg using this approach is consistent with the median LOD value of 0.020 ng/kg per congener (see Table 1) reported by the laboratories. The quantification limit L_Q equals $10S_\mathrm{Q}$ (where $S_\mathrm{Q} = 0.0055 + 0.0868L_\mathrm{Q}$), leading to a value of 0.417 ng/kg for LOQ. In fact, the standard deviation increases too sharply to estimate the LOQ by this approach. The ratio between LOQ/LOD increases from 3.04 (σ_B constant) to an unacceptable value of 19.75.

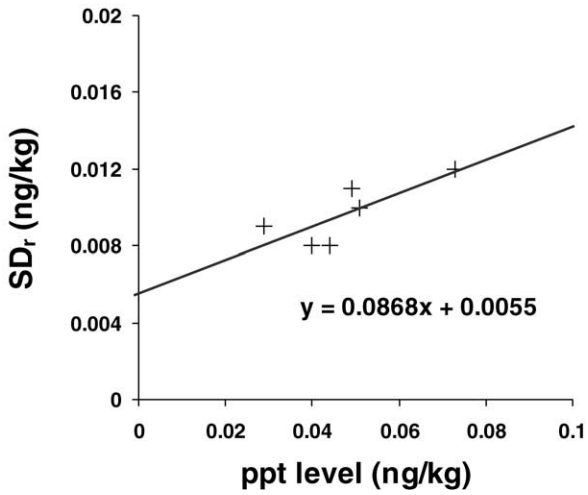

Fig. 2. Linear dependency of the standard deviation on PCDD/F congener levels.

2.4. The Eurachem approach for LOQ

The Eurachem approach [11] seems to be suitable to evaluate the LOQ from the results obtained in Part 1. The quantification limit here is defined on the lowest concentration of analyte that can be determined with an acceptable pre-established percentage of relative standard deviation. If a pre-established RSD_r of 20% is considered as an acceptable criteria for the ability to quantify dioxin congeners in feedingstuffs at sub-ppt level, then a corresponding LOQ of 0.049 ng/kg per congener is calculated by Eq. (1) (see also Fig. 1). The median LOQ for the less chlorinated congeners reported by laboratories (see Table 1) therefore corresponds to the analyte level for which an RSD_r of 20% can be achieved. Laboratories are thus able to provide reliable results at a level of 0.05 ng/kg per congener. Thus the precision criteria of Directive 2002/70/EC (i.e. CV < 15% at the maximum level) should probably be easily met for the confirmatory HRMS method.

2.5. S/N approach

The signal-to-noise ratio is still a widely used technique in gas chromatography. It is simple to implement, but the way of calculating the ratio, either manually or using commercial software, might lead to significant differences. Generally, LOD is calculated at *S*/*N* ratio of 3:1 and LOQ at *S*/*N* ratio of 10:1. For dioxin analysis, the Standard EN 1948-3 [6] does not make the distinction between detection limit and quantification limit. Only the concept of LOQ is defined at a *S*/*N* ratio of 3:1 with some specific requirements to be met. In order to evaluate the LOD/LOQ by the *S*/*N* approach, the results of the homogeneity test performed on 14 test portions (see Part 1) were exploited. The mean *S*/*N* ratios were calculated manually for the seven congeners below 0.11 ppt level. The mean values with their corresponding standard deviation were plotted against the ppt level, as can be seen in Fig. 3. The graph shows that the *S*/*N* ratio increases linearly with level, independently of the congener (i.e. tetra, penta and hexa-CDD/Fs). An estimation of the LOD/LOQ per congener using the regression line gives respectively 0.020 ng/kg at *S*/*N* of 3:1 and 0.061 ng/kg at *S*/*N* of 10:1. If the identification criteria, as retention time and isotope ratio, are fulfilled at *S*/*N* of 3:1, then the value of 0.020 ng/kg corresponds to an LOQ as defined in the EN 1948-3 [6].

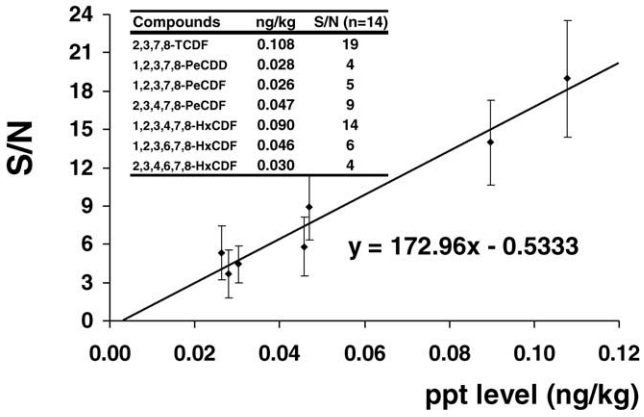

Fig. 3. Linear dependency of the *S*/*N* ratio on PCDD/F congener levels.

An overview of the different concepts and terminologies for detection and quantification limits for PCDD/F congeners (not present in blank procedure) in animal feed samples obtained by the different approaches is summarized graphically for the HRMS method in Fig. 4. The exploitation of the inter-laboratory data showed that a detection decision can already be taken at a level of 0.009 ng/kg per congener if a sample intake of roughly 30 g is used and if the blank

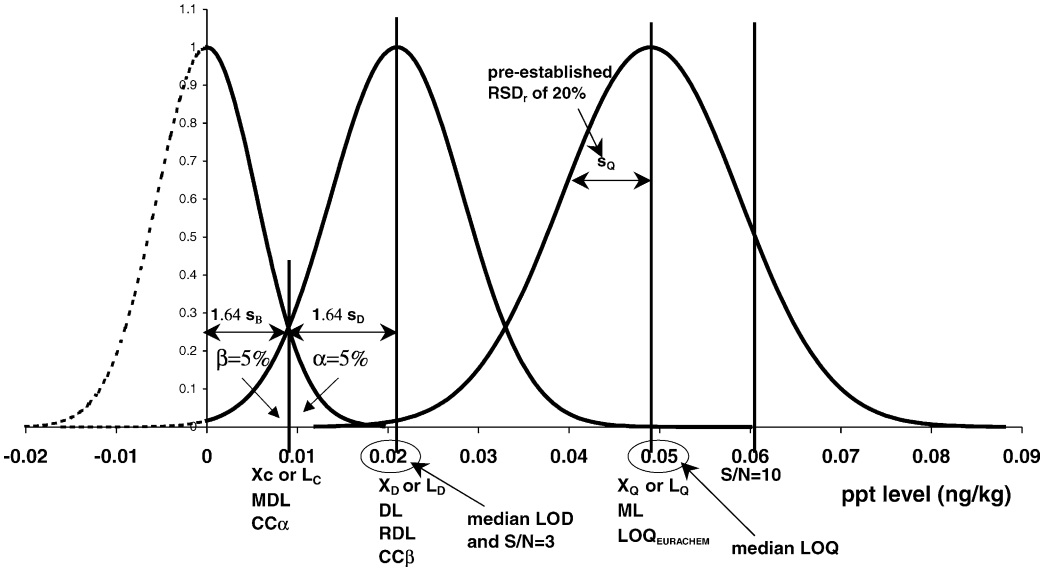

Fig. 4. Overview of the LOD/LOQ results for PCDD/F congeners in animal feed sample using the HRMS method.

procedure for the corresponding congener is 'really' blank. In these conditions, a detection limit beyond the detection decision at a level of approximately 0.02 ng/kg seems to be likely (i.e. $\alpha = 0.01\%$ and $\beta = 5\%$). Besides the detection limit, the statistical evaluation of data demonstrated that laboratories were able to provide reliable results at a quantification limit of 0.05 ng/kg with an accepted pre-established RSD_r of 20%.

3. Performances of the HRMS method in TEQ

Part 1 and the ensuing results presented here focused on the performances of the HRMS method for individual 2,3,7,8 PCDD/F congeners. However, legal limits in various food and feed matrices have been set in toxic equivalence of 2,3,7,8 TCDD in order to facilitate risk assessment. Taking this into account, the concept of toxic equivalency factors (TEFs) has been developed. The reporting of results and the interpretation against statutory limits therefore have to be discussed in those units. On the other hand, the analytical chemist performing dioxin analysis, using a specific congener technique like HRGC/HRMS method, can easily evaluate the method performances for each congener at different levels. Subsequently, the difficulty is demonstrating the method performances close to regulatory limits expressed in TEQ. This is not an easy task, as animal feedingstuffs samples generally contain a complex mixture of PCDD/F congeners. Moreover, patterns can vary from sample to sample depending on contamination sources. A combined total standard deviation in TEQ is then needed to assess both method performances and compliance to maximum limits. The combined total standard deviation in TEQ can be calculated by the square root of the sum of each congener's variance, also expressed in TEQ if each congener's determination can be seen as an independent random variable. For instance, the total PCDD/Fs reproducibility standard deviation (S_R) in TEQ is given by the following quadratic sum:

$$S_R = \sqrt{\sum_{i=1}^{17} (S_{Ri} \times TEF_i)^2} \quad (\text{ng WHO-TEQ/kg}) \qquad (4)$$

where S_{Ri} is the reproducibility standard deviation of the congener i (ng/kg) and TEF_i, its corresponding toxic equivalent factor.

The S_R value obtained in Part 1 for PCDD/Fs equals 0.15 ng WHO-TEQ/kg, using each congener inter-laboratory reproducibility standard deviation calculated by ISO 5725 (see Table 4, Part 1). By comparison, the inter-laboratory precision model, represented here by Eq. (2), can be reorganized to give

$$S_{Ri} = 0.0169 + 0.119 c_i \quad (\text{ng/kg}) \qquad (5)$$

Thus, the total S_R in TEQ recalculated using the precision model (Eq. (5)) gives $S_R = 0.18$ ng WHO-TEQ/kg, indicating that the model fits well with the data. Individual congener precision, $S_{Ri} = f(c_i)$, can therefore be used to assess the precision of the HRMS method in TEQ close to maximum limits. Unfortunately, as already mentioned, the correlation between individual congener precision and TEQ precision, using the model, can only be done if the congener pattern is known. Since congener patterns in animal feedingstuffs mainly depend on environmental contamination sources, many different PCDD/F patterns exist [13]. Nevertheless, with the aim of assessing the HRMS method's precision in TEQ close to regulatory limits, an example of a simple homogeneous pattern was selected. At 0.75 and 0.5 ng WHO-TEQ/kg, it corresponds respectively to an individual congener level of 0.221 and 0.148 ng/kg. Then, using Eq. (5) to calculate their corresponding S_{Ri} and Eq. (4) to obtain the combined total S_R in TEQ, an assessment of the inter-laboratory reproducibility precision in TEQ is presented in Fig. 5. At the LOQ, the S_R in TEQ was calculated according to the median values reported in Table 1. Assuming the normality distribution at the different levels investigated, S_R represents the Gaussian distribution width. The figure shows that the action level and maximum level are significantly different in terms of reproducibility precision: the data distributions partially overlap by less than 2% and their inter-laboratory reproducibility CVs are respectively 10.5 and 8.8%. The study also demonstrated that with the current state-of-the-art technology for dioxin analysis using HRMS technique, the lowest reliable result for quantification was 0.17 ng WHO-TEQ/kg. The future establishment of target values for PCDD/Fs in animal feed samples surely needs to take into account this analytical recommendation. These analytical considerations are of primary importance and have been highlighted by recent conclusions drawn from experimental studies investigating the carry-over rates of dioxin from contaminated feed and soil to laying hens' eggs. Traag et al. [14] concluded that feeding laying hens for 56 days with contaminated feedingstuffs at 0.75 ng WHO-TEQ/kg can lead to exceeding the maximum level of 3 pg WHO-TEQ/g fat in eggs. The EU action limit of 2 pg WHO-TEQ/g fat in eggs might even have already been exceeded with contaminated feedingstuffs at 0.2 ng WHO-TEQ/kg. Thus, continuous efforts to develop analytical methods to reliably achieve the lowest detection limit are more than ever topical.

According to the Decision 2002/657/EC [8], normally for the determination of group B substances with permitted limit (PL) or maximum level, the decision limit (CC_α) and the detection capability (CC_β) have nothing in common with detection and quantification limits. CC_α and CC_β are only related to LOD/LOQ or to analytical performances of the method close to detection for substances of group A with no permitted limit. In spite of this, these new analytical aspects were introduced here for detection and quantification limit (see Fig. 4) with the aim of presenting an overview of the different concepts and to do the relationship between the terminologies but also because dioxin levels in food and feed

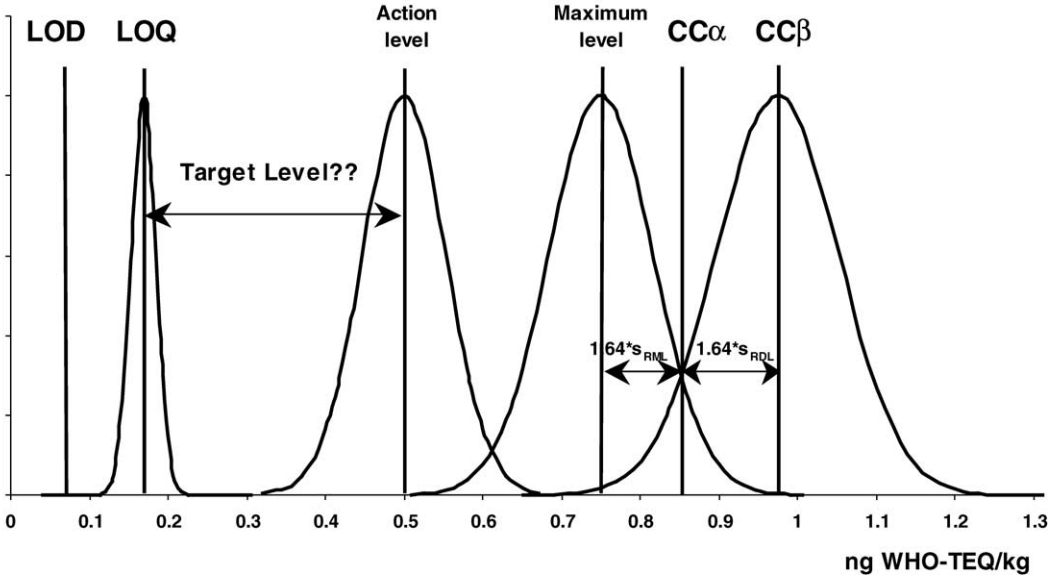

Fig. 5. Precision of the HRMS method in TEQ units for dioxin in feedingstuffs at different levels of interest.

are generally characterized by background contaminated levels. Even with PL, for those substances, the approach by the HRMS method is different from the classical residue approach described in the Decision 2002/657/EC because CC_α and CC_β definitions can be used here in both cases. On one hand, the information for compliance assessment is needed close to maximum level (using CC_α and/or CC_β) and on the other hand, relevant information on analytical performances close to LOD/LOQ (or CC_α/CC_β or others) is also needed for risk and dietary intake assessments.

In the case of B substances, CC_α is the level above which it can be decided with a statistical certainty of $1 - \alpha$ that the PL has been truly exceeded [15]. A decision of non-compliance above CC_α hence becomes quite likely, and above CC_β, extremely likely. The inter-laboratory reproducibility standard deviation in TEQ S_{RML} (see Fig. 5) calculated by the model at the maximum limit can be used to evaluate CC_α. CC_α is then obtained by multiplying S_{RML} by 1.64 (one-tailed value of the distribution) at a 95% confidence level to give a value of 0.86 ng WHO-TEQ/kg. Identically, CC_β can be assessed by multiplying S_{RDL}, recalculated at the decision limit (CC_α), by 1.64, yielding a value of 0.97 ng WHO-TEQ/kg. These values can then be used to facilitate the interpretation and/or the decision-making for compliance or non-compliance of a lot/sublot against statutory limit. In our example, the HRMS method precision study permitted assessing an indicative CC_α value of 0.86 ng WHO-TEQ/kg, above which the rate of false non-compliant decisions is reduced to less than 5%. If a rate of 5% of false compliance is required for decision-making, then the lot is not accepted if the result is above the indicative detection capability value of 0.97 ng WHO-TEQ/kg. Note that at CC_β, the rate of false non-compliant decisions is reduced to 0.03%.

4. Measurement uncertainty

Another way of reporting and interpreting analytical results for compliance to legal limits is to estimate measurement uncertainty. An analytical result takes the form of 'a ± U' where the expanded uncertainty U provides an interval within which the true value is believed to lie at a set level of confidence. There are numerous ways and procedures, depending on the data available, to estimate an uncertainty budget. The use of collaborative trial data, e.g., according to the Standard ISO 5725, is one possibility [16]. The inter-laboratory reproducibility standard deviation S_R may cover many uncertainty sources, but it is also necessary to identify other significant sources that may be not covered by the study (such as sampling, sample pre-treatment, method bias, variation in conditions, etc.). These possible identified sources are not considered as significant if they are less than one-third of S_R. On the other hand, it is not an easy task to assess these components and to express them as a standard deviation. Regarding the sampling, only the sub-sampling used to perform analyses can be assessed. As described in Part 1, the participants were asked to use a minimum of 20 g, and preferably 30 g, of sample intakes out of a 100 g umber glass bottle to perform the analysis. All the laboratories respected this request. The homogeneity test carried out in Part 1 on 30 g of sample intakes provided a 'sampling standard deviation' for each congener, and is assumed to represent only a component of the sampling uncertainty source. The global uncertainty due to sampling for an independent test analysis on a raw animal feedingstuffs is more difficult to evaluate. The measurement is focused here on the analytical part. For the other possible sources, the participants did not report any specific pre-treatment of the sample. Method bias is difficult to evaluate since, until to-

Table 2
Contribution of the uncertainty sub-sampling S_{sam} to S_R

Compounds	S_R	S_{sam}	Combined uncertainty S
2,3,7,8-TCDD	0.0096*	0.0014	0.0097
1,2,3,7,8-PeCDD	0.0194		0.0194
1,2,3,4,7,8-HxCDD	0.0414	0.0038	0.0415
1,2,3,6,7,8-HxCDD	0.2165	0.0415	0.2204
1,2,3,7,8,9-HxCDD	0.1036	0.0257	0.1068
1,2,3,4,6,7,8-HpCDD	14.5087	1.4018	14.5763
1,2,3,4,6,7,8,9-OCDD	109.0537	18.8508	110.6709
2,3,7,8-TCDF	0.0269	0.0075	0.0279
1,2,3,7,8-PeCDF	0.0152		0.0152
2,3,4,7,8-PeCDF	0.0208	0.0029	0.0210
1,2,3,4,7,8-HxCDF	0.0283		0.0283
1,2,3,6,7,8-HxCDF	0.0115	0.0022	0.0117
2,3,4,6,7,8-HxCDF	0.0116		0.0116
1,2,3,7,8,9-HxCDF			0.0000
1,2,3,4,6,7,8-HpCDF	0.3333		0.3333
1,2,3,4,7,8,9-HpCDF	0.0556		0.0556
1,2,3,4,6,7,8,9-OCDF	2.4297		2.4297
Total PCDD/Fs in TEQ	0.1496	0.0151	0.1503

* Cofino's standard deviation.

day, no reference material has been available. The variations in conditions were included in S_R since no specific requirements were enforced for extraction and clean up (see Part 1). Only the HRMS using the isotopic dilution technique was imposed. Table 2 shows, according to the homogeneity test performed in Part 1, that the sub-sampling source characterized here by S_{sam} is not only negligible for all the PCDD/F congeners, but also in TEQ compared to S_R. Note that the homogeneity test did not provide a sampling standard deviation for all the congeners (see Part 1). The in-house studies on the same material, based on a long-term precision study (i.e. intra-laboratory reproducibility) and used as an internal quality control, confirmed a standard deviation close to S_R in TEQ. For instance, for a period covering the years 2002 and 2003, the intra-laboratory standard deviation ($n = 121$) on the same material was equal to 0.19 ng WHO-TEQ/kg, compared to 0.15 ng WHO-TEQ/kg obtained for the present study. Note that the intra-laboratory reproducibility standard deviation should in general be lower than the inter-laboratory standard deviation. Here, this situation can be explained by a restricted number of expert laboratories involved in the inter-laboratory study. If we can reasonably consider that S_R represents the uncertainty associated with the result, then the precision model (Eq. (5)) can be applied to recalculate the corresponding uncertainty associated with each congener amount. Thus, each analyte above LOQ, with its corresponding uncertainty, can be reported. The present study therefore provides an indication of expected uncertainty for the HRMS method. Obviously, each laboratory has to correctly evaluate its own uncertainty corresponding to its analytical method.

One question directly arises on the method of reporting dioxin results from monitoring programs. Indeed, more than 95% of results, especially in animal feedingstuffs control, are background contamination, i.e., most of the congeners are not detected. By applying the lower bound approach, the sum of each congener contribution is zero, but their uncertainty is not zero. It is widely accepted that below LOQ, the method's performance becomes insufficient for acceptable quantitation. In that region of 'high uncertainty', as the congener level drops, the relative uncertainty associated with the result tends to increase and may become larger compared to the result. For the present survey, an inter-laboratory reproducibility relative standard deviation of 46% is already achieved at the LOQ and rises to roughly 100% at the LOD. So the associated uncertainty of the observation in this region is certainly more appropriately represented by a rectangular distribution focused around the LOD. Indeed, there is reason to believe that extreme values are likely. The upper limit of the interval is set at the LOQ, while the lower limit falls below zero, as can be seen in Fig. 6. In

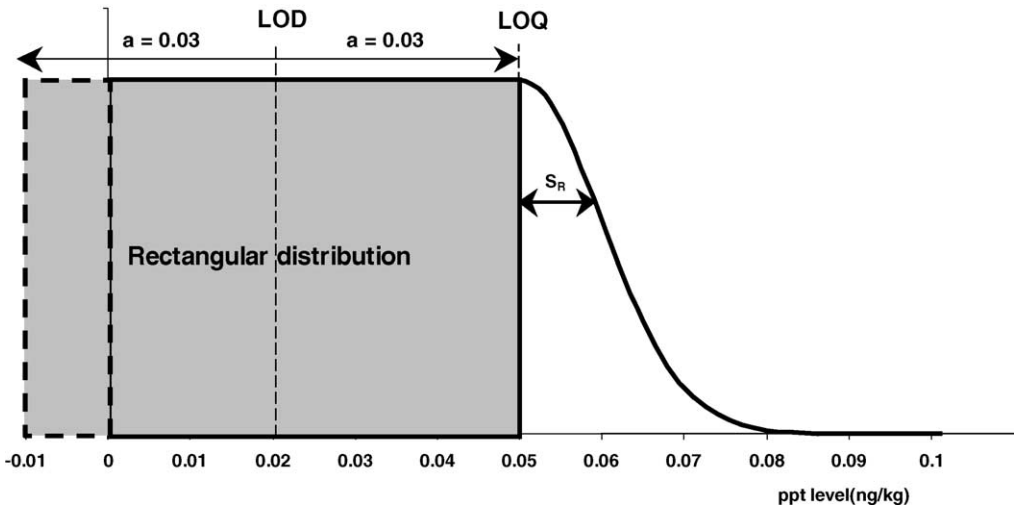

Fig. 6. Uncertainty estimation below LOQ.

this region, negative observations are quite possible, even though the true value of the measurand cannot be (e.g. when the value of an analytical blank is subsequently used to correct results close to LOQ). The rectangular distribution is hence characterized by limits of ±0.03 ng/kg and the corresponding standard deviation equals $0.03/\sqrt{3}$ ng/kg.

5. Reporting of results and interpretation

Let us take a real animal feed sample containing a sum of PCDD/Fs close to the maximum level. Table 3 summarizes the individual congener levels found (in ppt). Among these, three congeners were not detected (i.e. 2,3,7,8 TCDD; 1,2,3,7,8 PeCDD and 1,2,3,7,8,9 HxCDF). The congeners' LOQs in TEQ used here are the median values already reported previously (see Table 1). The recommendation for dioxin reporting [1] indicates that results have to be reported in lower, middle and upper bound levels. The results for the sum of PCDD/Fs in TEQ are respectively 0.78, 0.84 and 0.89 ng WHO-TEQ/kg. The uncertainty associated with the result needs to be taken into account when assessing compliance. Then, the corresponding estimated uncertainties were calculated for each congener at or above LOQ using the precision model S_R (Eq. (5)). For the three non-detected congeners, their uncertainty was evaluated by using the standard deviation of the rectangular distribution ($0.03/\sqrt{3}$ ng/kg) for lower and middle bound approaches. The total uncertainty associated with the sum in TEQ was then calculated using the law of propagation of uncertainties by applying the quadratic sum (Eq. (4)). The combined uncertainties in TEQ are respectively 0.060 ng WHO-TEQ/kg for lower and middle bound, and 0.063 ng WHO-TEQ/kg for upper bound. The final result should take the form of 'a ± U', where U is calculated by multiplying the combined uncertainty by a coverage factor of 2 at a 95% confidence level. Fig. 7 graphically represents the three situations with their associated uncertainties. Whatever the approach used, results are above the maximum limits. However, the maximum limit lies in the range of uncertainties quoted for two of them. Non-compliance is not demonstrated beyond reasonable doubt for the lower bound and middle bound results. A recent SANCO working document [17], dealing with the reporting of results and interpretation against statutory limits, recommended that measurement uncertainty has to be estimated and reported at levels above a legislation limit. Thus, only if the reported level minus the expanded U uncertainty is greater than the maximum limit can one be led to conclude beyond reasonable doubt that the sample is non-compliant. In our example, the result expressed according to the upper bound approach (i.e. 0.89 ng WHO-TEQ/kg) fulfills the requirement and can therefore be considered as non-compliant beyond reasonable doubt. Thus, the European legislation clearly pointed out that the decision regarding compliance or non-compliance of a lot against maximum limit does not favor consumer protection. Instead of reporting a result with

Table 3
Reporting of results

Congeners	WHO-TEF	Level (ng/Kg)	Lower bound (ng WHO-TEQ/kg)	Middle bound (ng WHO-TEQ/kg)	Upper bound (ng WHO-TEQ/kg)	LOQ (ng WHO-TEQ/kg)	Lower bound S_R	Middle bound S_R	Upper bound S_R
2,3,7,8 TCDD	1	nd	nd	0.025	0.0500	0.0500	0.0173	0.0173	0.0228
1,2,3,7,8 PeCDD	1	nd	nd	0.025	0.0500	0.0500	0.0173	0.0173	0.0228
1,2,3,4,7,8 HxCDD	0.1	0.069	0.007	0.007	0.007	0.0050	0.0251	0.0251	0.0251
1,2,3,6,7,8 HxCDD	0.1	0.840	0.084	0.084	0.084	0.0050	0.1165	0.1165	0.1165
1,2,3,7,8,9 HxCDD	0.1	0.352	0.035	0.035	0.035	0.0050	0.0587	0.0587	0.0587
1,2,3,4,6,7,8 HpCDD	0.01	38.887	0.389	0.389	0.389	0.0036	4.6289	4.6289	4.6289
OCDD	0.0001	316.115	0.032	0.032	0.032	0.000032	37.5081	37.5081	37.5081
2,3,7,8 TCDF	0.1	0.448	0.045	0.045	0.045	0.0050	0.0701	0.0701	0.0701
1,2,3,7,8 PeCDF	0.05	0.103	0.005	0.005	0.005	0.0025	0.0291	0.0291	0.0291
2,3,4,7,8 PeCDF	0.5	0.254	0.127	0.127	0.127	0.0250	0.0471	0.0471	0.0471
1,2,3,4,7,8 HxCDF	0.1	0.168	0.017	0.017	0.017	0.0050	0.0368	0.0368	0.0368
1,2,3,6,7,8 HxCDF	0.1	0.130	0.013	0.013	0.013	0.0050	0.0323	0.0323	0.0323
1,2,3,7,8,9 HxCDF	0.1	nd	nd	0.003	0.005	0.0050	0.0173	0.0173	0.0228
2,3,4,6,7,8 HxCDF	0.1	0.089	0.009	0.009	0.009	0.0050	0.0274	0.0274	0.0274
1,2,3,4,6,7,8 HpCDF	0.01	1.812	0.018	0.018	0.018	0.0006	0.2318	0.2318	0.2318
1,2,3,4,7,8,9 HpCDF	0.01	0.175	0.002	0.002	0.002	0.0008	0.0376	0.0376	0.0376
OCDF	0.0001	7.118	0.001	0.001	0.001	0.000018	0.8611	0.8611	0.8611
Total WHO-TEQ-PCDD/F		366.56	0.78	0.84	0.89	0.17	0.060	0.060	0.063

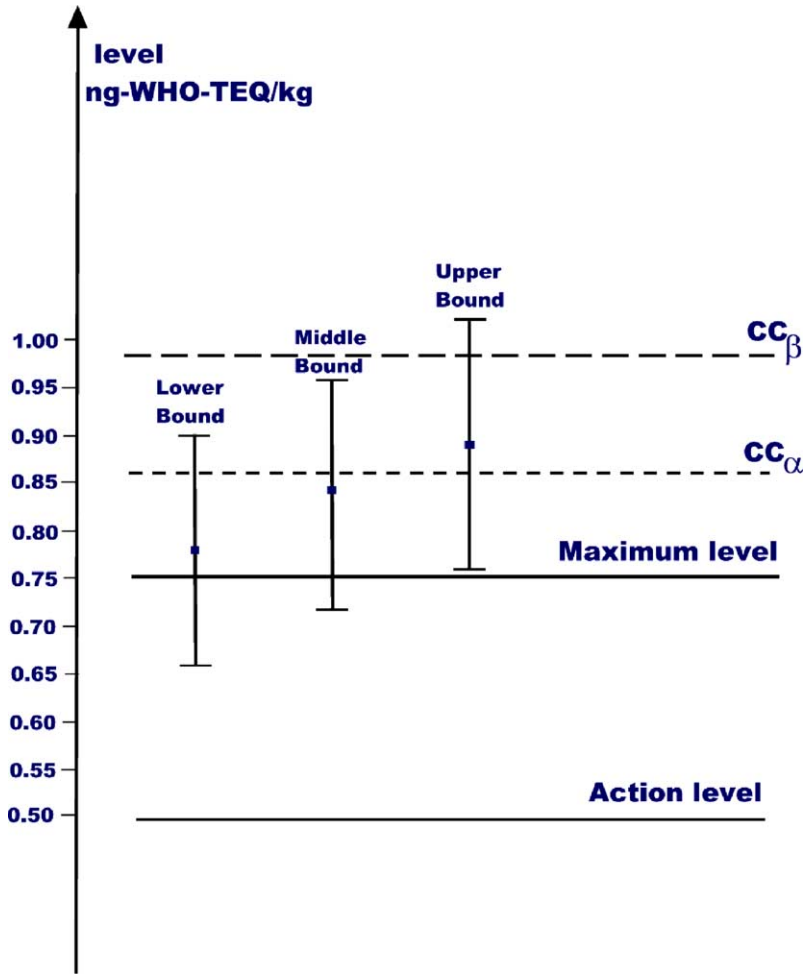

Fig. 7. Graphical representation of reported results.

its associated uncertainty to demonstrate compliance with maximum limit, the other approach, based on the Commission Decision 2002/657/EC [8] for substances for which permitted limits were established, is to shift the decision limit beyond the maximum level at CC_α or even CC_β. In our case, the indicative CC_α and CC_β limits calculated previously are represented on Fig. 7. If a non-compliance decision is taken at CC_α, then the reported results lead to the same conclusion as the uncertainty approach, meaning that the upper bound result is non-compliant. On the other hand, if the decision is now taken at CC_β, then all the results are compliant.

6. Conclusions

The statistical treatment of data from an inter-laboratory study and the exploitation of these results enable us to partially answer the open question: does the HRMS method comply with the European pro-active approach for dioxin analysis in animal feedingstuffs? Indeed, the results have shown, with the state-of-the-art HRMS method that reliable results can be easily provided up to a value of 0.17 ng WHO-TEQ/kg. This is completely in agreement with the current European Council Directive 2001/102/CE. However, the pro-active approach is to decrease regulatory levels, whereas the number of chemical compounds in TEQ will increase. This might then lead to possible problems in assessing compliance of a product against lessening statutory limits. On the other hand, the requirements for such residues in animal feedingstuffs, placed at the top of the food chain, have to be as strict as possible in order to ensure compliance with limits in foodstuffs. The objective of this paper was also to attract attention to analytical considerations based on the performances and limitations of the reference method for the future establishment of target levels.

A non-exhaustive list of definitions and concepts for LOD/LOQ determination has been introduced and discussed. For this study, the results showed good agreement between the different approaches used, but nevertheless there is clearly a need to harmonize these concepts.

Based on the inter-laboratory data, a precision model for repeatability and reproducibility covering several orders of magnitude in the range of ppt level for PCDD/F congeners has been developed. The model can certainly provide helpful information and indicative values of standard deviation that can be expected by laboratories developing their own analytical method.

Finally, the reporting of results and the assessment for compliance against limits using two different approaches have been broached. One is to report a result close to a maximum level with its uncertainty measurement. The decision of non-compliance beyond reasonable doubt is only taken if the result minus its expanded uncertainty is above the maximum limit. The other is to report only the result and to shift the decision level above the maximum level, where it can be decided with statistical certainty that the maximum level has been truly exceeded. Both approaches favor producers rather than consumer protection.

Acknowledgements

The authors thank Ir. J. Widart for fruitful discussion. Financial support of the Fonds Européen de Développement Régional (FEDER), the Fonds Social Européen (FSE) and partially by the Service Public Fédéral Santé Publique (Formerly Ministère des Classes Moyennes et de l'Agriculture DG6) is also acknowledged.

References

[1] JECFA recommendations, Joint FAO/WHO Expert Committee on Food Additive, in: Proceedings of the 57th Meeting, WHO Food Additives Series 48.
[2] L.A. Currie, Anal. Chim. Acta 391 (1999) 127.
[3] L.A. Currie, Pure Appl. Chem. 67 (1995) 1699.
[4] ISO 11843-1, Capability of detection: Part1: Terms and Definitions, International Standard Organisation, Geneva, 1997.
[5] Proceedings of the United Nations Environment Programme, Chemicals, Geneva, Switzerland, 2003, p. 24.
[6] European Standard EN 1948-3 (1996).
[7] EPA method 1613 revision B, Tetra-through Octa-Chlorinated Dioxins and Furans by Isotope Dilution HRGC/HRMS, October, 1994.
[8] Commission Decision 2002/657/EC, Off. J. Eur. Commun. L 221 (August 12) (2002) 8.
[9] L.H.Keith, Environmental Sampling and Analysis, a Practical Guide, CRC, 1991.
[10] Council Directive 96/23/EC, Off. J. Eur. Commun. L 125 (May 23) (1996) 10.
[11] Eurachem guide: the fitness for purpose of analytical methods. A Laboratory Guide to Method Validation and Related Topics, LGC, Teddington UK, 1998.
[12] J. Ferrario, C. Byrne Jr., A.E. Dupuy, Chemosphere 34 (1997) 2451.
[13] H. Hagenmaier, C. Lindig, J. She, Chemosphere 29 (1994) 2163.
[14] W. Traag, K. Kan, M. Zeilmaker, R. Hoogenboom, Organohalogen Compd. 61 (2003) 381.
[15] K. DeWasch, H.F. DeBrabander, D. Courtheyn, N. Van Hoof, S. Poelmans, H. Noppe, Eur. Food Chem. XII 1 (2003) 45.
[16] S.L.R. Ellison, M. Rosslein, A. Williams, Eurachem/CITAC guide, Quantifying Uncertainty in Analytical Measurement, 2nd ed., 2000.
[17] Working Document SANCO/0064/2003-rev. 2.