THE DIOXIN FUNCTION AS A USEFUL TOOL FOR ASSESSING DIOXIN LABORATORY PERFORMANCES IN PROFICIENCY TESTS

Eppe G[1], Van Cleuvenbergen R[2], Smastuen Haug L[3], Becher G[3], De Pauw E[1]

1 CART Mass Spectrometry Laboratory, Chemistry Department, University of Liège, Allée de la Chimie 3, B-6c Sart-Tilman, B-4000 Liège, Belgium
2 VITO, Flemish Institute for Technological Research, Boeretang 200, B-2400 Mol, Belgium
3 Department of Analytical Chemistry, Division of Environmental Chemistry, Norwegian Institute of Public Health, P.O. Box 4404 Nydalen, 0403 Oslo, Norway

Abstract

This study reports on precision characteristics achieved by the gas chromatography-isotope dilution-high resolution mass spectrometry (GC-ID-HRMS) reference method for polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and dioxin-like polychlorinated biphenyls (DL-PCBs) in food and feed in two interlaboratory method-performance studies among expert laboratories with long-standing experience in this field. Striking linear functions in log scale between reproducibility standard deviation and congener's level over a concentration range of 10^{-8} to 10^{-14} g g^{-1} fresh weight are observed. The data fit very well to a Horwitz-type function of the form $s_R = 0.153c^{0.904}$, where s_R and c are dimensionless mass ratios expressed in pg g^{-1} on fresh weight, regardless of the nature of the toxic congeners, food and feed matrices, or sample preparation methods. We demonstrate that the proposed function is suitable for use as a fitness-for-purpose criterion for proficiency testing (PT) or interlaboratory comparisons on PCDDs, PCDFs and DL-PCBs compounds in food.

Introduction

Participation in proficiency testing (PT) exercises is a useful way for assessing the quality of a laboratory's results. In chemical analysis, the statistical evaluation of PT results is based on the scoring system as recommended in the International Harmonized Protocol[1]. In this system, the participant's result is converted into a z-score that gives a valuable indication of the performance of the laboratory. In fact, a z-value is nothing more than the estimate of the error in the result scaled in standard deviation units (σ_p). However, to assess correctly the performances of laboratories we need to define target standard deviation in such a way that it should describe the end-user's requirements. In addition, it is of primary importance for dioxin laboratories to assess their performances not only by a single result expressed in TEQ but also for all the congeners that significantly contribute to the TEQ value.

The main features that characterize a PCDD/F measurement by GC-ID-HRMS are the remarkable precision and trueness achieved in a concentration range that is currently not explored by any other applications in chemical analysis in the food sector. With the DL-PCBs, the measurement of the 29 toxic congeners in a biological sample can cover several orders of magnitude from sub-parts-per-trillion to parts-per-billion. Using precision models described in the literature[2,3], it is difficult to assess whether the reproducibility obtained is fit for purpose. The aim of this paper is to provide an empirical relationship between s_R and the 2,3,7,8-PCDD/F and DL-PCB congener level in food and feed based on data collected from collaborative trials involving a restricted number of expert laboratories. The proposed model is suggested as suitable for use as a fitness-for-purpose criterion for PCDDs, PCDFs and DL-PCBs in PT exercises. Application of the proposed function to annually performed international PT schemes on dioxins in food is discussed.

Materials and Methods

Interlaboratory studies : EU Feed 2001, DIFFERENCE 2004

The EU Feed 2001 study design and the statistical treatments of data have been described elsewhere[4]. Briefly, the statistical data treatment of the interlaboratory comparison was carried out according to the Standard ISO 5725-2. It is based on the classical ANOVA technique that gives an estimation of the gross average, intra-laboratory and inter-laboratory variances, repeatability and reproducibility of the method.

A second interlaboratory exercise took place in 2004 in the framework of the European project called DIFFERENCE. It concerned a feasibility study on five candidate CRMs in food and feed. The materials were herring tissue, pork meat, whole milk, herring-oil, and pig feed material, respectively. Ten expert GC-ID-HRMS laboratories were invited to participate in the certification exercise; most of them had also been involved in the EU Feed 2001 study. Details on the statistical treatments of data and the technical report are available elsewhere[5,6].

Proficiency Tests: Dioxins in FOOD 2004-2006

The fifth, sixth and seventh rounds of Interlaboratory Comparison on Dioxins in Food was conducted by the Norwegian Institute of Public Health on the determination of seventeen 2,3,7,8-substituted PCDD/Fs, DL-PCBs as well as six marker PCBs (except 2004) and eight polybrominated diphenylethers (PBDEs). The objectives were to offer a quality assurance instrument for the participating laboratories, to assess the between laboratory reproducibility, as well as the readiness of expert laboratories world-wide to determine levels of chlorinated and brominated persistent organic pollutants in regular foodstuffs. The study was performed on sample homogenates of chicken meat, trout, palm oil, reindeer meat, herring filet, cod liver oil, egg yolk, halibut filet and breast milk. The consensus concentration for each analyte in the nine food samples was determined as follows: The median of all reported concentrations for each analyte was calculated. All values above two times the median were then removed from the calculation. The consensus median was calculated from the remaining data.

Results and discussion

The outcome of interlaboratory method performance studies for dioxin-like compounds in food and feed allowed for modeling the precision of the state-of-the-art GC-ID-HRMS method by an empirical Horwitz-type function, termed Dioxin function:

$$s_R = 0.153c^{0.904} \quad \text{if } 0.01 \text{ pg g}^{-1} \leq c \leq 10000 \text{ pg g}^{-1} \quad (\text{i.e. } 10^{-14} \leq c \leq 10^{-8}) \quad (1)$$

Where c is the PCDD/Fs and DL-PCBs analyte concentration in pg g^{-1} fw mass ratio.

A complete description of its establishment can be found elsewhere[7]. The results fell into a surprisingly uniform pattern regardless of the nature of the toxic congeners, food and feed matrices and the slight differences in sample preparation protocols. Hence, s_R takes account of the following variations:

- Variation within runs in a laboratory
- Different runs (time, analysts, calibration) within a laboratory
- The differences in the method protocols between laboratories
- Many other systematic errors of individual laboratories

An interesting application of the dioxin function might be its use as a target standard deviation in performance assessment of PT exercises. According to the International Harmonized Protocol[1], the z-scores are calculated as follows:

$$z = \frac{x - X}{\sigma_p} \quad (2)$$

Where x: lab result; X: assigned value, σ_p target value for the standard deviation.
One of the main features that characterize a PCDD/Fs and DL-PCBs measurement in a biological sample is the wide range of contamination levels between congeners. As an example, Table 1 shows the consensus medians in pg g^{-1} fresh weight of the herring tissue material from 'Dioxins in FOOD 2005'. Five orders of magnitude separate the 1,2,3,7,8,9-HxCDF level from PCB 118 level. It is clear that the analytical requirement to assess their individual performances need to be adjusted accordingly.

Table 1: Results from Dioxins in FOOD 2005, Herring tissue

Congeners	consensus median (pg g^{-1})	standard deviation (pg g^{-1})	S_R (pg g^{-1})*	ratio (SD/S_R)
2,3,7,8 TCDD	0,097	0,023	0,019	1,24
1,2,3,7,8 PeCDD	0,170	0,049	0,031	1,59
1,2,3,4,7,8 HxCDD	0,040	0,012	0,008	1,44
1,2,3,6,7,8 HxCDD	0,116	0,032	0,022	1,46
1,2,3,7,8,9 HxCDD	0,029	0,009	0,006	1,47
1,2,3,4,6,7,8 HpCDD	0,070	0,022	0,014	1,62
OCDD	0,112	0,046	0,021	2,18
2,3,7,8 TCDF	1,634	0,330	0,238	1,38
1,2,3,7,8 PeCDF	0,239	0,044	0,042	1,05
2,3,4,7,8 PeCDF	0,642	0,107	0,102	1,05
1,2,3,4,7,8 HxCDF	0,097	0,029	0,019	1,56
1,2,3,6,7,8 HxCDF	0,070	0,021	0,014	1,54
2,3,4,6,7,8 HxCDF	0,081	0,026	0,016	1,64
1,2,3,7,8,9 HxCDF	0,010	0,005	0,002	2,04
1,2,3,4,6,7,8 HpCDF	0,050	0,017	0,010	1,66
1,2,3,4,7,8,9 HpCDF	0,010	0,008	0,002	3,18
OCDF	0,029	0,017	0,006	2,72
PCB 77	26	6,6	2,9	2,26
PCB 126	6,8	1,5	0,9	1,71
PCB 169	1,4	0,36	0,2	1,75
PCB 81	1,1	0,35	0,2	2,13
PCB 105	297	72	26,3	2,75
PCB 114	11	4,3	1,4	3,15
PCB 118	1024	280	80,5	3,48
PCB 123	12	5,4	1,5	3,64
PCB 156	99	21	9,7	2,14
PCB 157	28	6,3	3,1	1,99
PCB 167	79	18	8,0	2,23
PCB 189	12	2,7	1,4	1,89

*: calculated with equation (1)

In Table 1, we also reported the standard deviation (SD) from the NIPH report. SDs were calculated after removing obvious outliers. In addition, for each consensus median, the target s_R was calculated via equation (1). Finally, the ratio between SD and s_R is calculated. Proficiency tests 'Dioxin in FOOD' is a

worldwide PT currently comprising about 100 laboratories. These exercises are real life studies comprising a broad range of analytical laboratories (as opposed to the especially designed studies, for example, to establish the dioxin function), and there may be additional sources of errors. A SD to s_R ratio >1 represents these additional effects. It is relatively reproducible among PCDD/F congeners and also for the DL-PCBs. However, some congeners 1,2,3,7,8,9-HxCDF, 1,2,3,4,7,8,9-HpCDF, OCDF (and in some cases OCDD) do not behave in the same way. These congeners are generally not detected by more than 50% of the laboratories and have to be treated separately. It is however doubtful if it is necessary to assess performances on congeners generally close to detection limits and insignificantly contributing to the TEQ. We did the same exercise for the nine materials of the last three years PT exercises and calculated the median ratio separately for the PCDD/Fs as well as for the DL-PCBs. Table 2 summarizes the results representing more than 10000individual results reported by the participants.

Table 2: Overview of the median ratio between SD and s_R for PCDD/Fs and DL-PCBs

		PCDD/Fs		DL-PCBs	
	samples	Median ratio (SD/S_R)	RSD (%) of ratio	Median ratio (SD/S_R)	RSD (%) of ratio
FOOD 2004	chicken	2,0	20	2,8	17
	trout	1,9	16	3,4	22
	palm oil	2,1	19	*	*
FOOD 2005	reindeer	1,8	12	2,6	23
	cod liver oil	1,9	15	2,8	31
	herring	1,6	20	2,4	27
FOOD 2006	egg yolk	1,9	20	2,2	25
	halibut	1,7	17	2,7	21
	breast milk	1,4	12	1,8	35

*: very poor precision was reported for DL-PCBs in palm oil (60-70% RSD), related to extremely low levels of DL-PCBs

For PCDD/Fs, we observe a slight improvement for the last two exercises compared to the 2004 exercise. The median ratio was around 2.0 for FOOD 2004 and it drops just under 2.0 for FOOD 2005 and 2006. The same trend can also be highlighted for DL-PCBs. Based on these results, we propose to use the dioxin function as target standard deviation for individual congeners z-scores assessment and to interpret the scoring system according to the following rate: $z \leq \pm 2$ (satisfactory); $\pm 2 < z < \pm 3$ (questionable); $z \geq \pm 3$ (unsatisfactory).

References

1. Thompson M., Ellison S.L.R., Wood R., *Pure Appl. Chem*.; 2006; 78; 145:196.
2. Horwitz W., Kamps L.R., Boyer, K.W., *J. Assoc. Off. Anal. Chem*.; 1980; 63; 1344:1354.
3. Thompson M., Analyst ; 2000; 125; 385:386
4. Eppe G., Cofino W.P., De Pauw E., *Anal. Chim. Acta*.; 2004; 519; 231:242.
5. Van Cleuvenbergen R., Final report WP4 DIFFERENCE project (G6RD-CT-2001-00623); 2005.
6. van Leeuwen S.P.J., Van Cleuvenbergen R., Abalos M., Pasini A.-L., Eriksson U., Cleemann M., Hasjlova J., de Boer J., *Trends Anal. Chem*.; 2006; 25; 397:409.
7. Eppe G., Van Cleuvenbergen R., Smastuen Haug L., Boulanger B., Becher G., De Pauw E., *Chemosphere* ; 2007, accepted.