QUALITY AND UNCERTAINTY IN THE ANALYSIS OF POPs IN FOOD - EXPERIENCES FROM 10 YEARS OF WORLD-WIDE INTERLABORATORY COMPARISON

Becher G, Haug LS, Broadwell SL, Bruun Bremnes NM*

Division of Environmental Medicine, Norwegian Institute of Public Health, P.O.Box 4404 Nydalen, 0403 Oslo, Norway

Introduction

Persistent organic pollutants (POPs), such as polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDDs/PCDFs), dioxin-like polychlorinated biphenyls (dl-PCBs) and brominated flame retardants (BFRs) are globally distributed in practically all environmental compartments and pose a risk of causing adverse effects to human health and the environment. Humans are exposed to these POPs mainly through the diet with food of animal origin usually being the predominant source. There is considerable public, political and scientific concern regarding the adverse effects on human health of long-term exposure to even minute amounts of dioxin-like compounds.

In order to protect humans from dietary exposure to dioxin-like compounds, maximum, action and target concentrations in food are set in the European Union as well as in other countries. Recent incidents with dioxin contamination in the food chain have demonstrated the necessity of frequent monitoring of the presence of dioxin-like compounds in foodstuffs. Similarly, the European Food Safety Authority (EFSA) has encouraged member countries to monitor BFRs in food in order to create a sufficient data base for risk assessment. In order to ensure sufficient analytical quality, official control laboratories within the EU as well as all analytical laboratories that are accredited according to ISO/IEC 17025:2005 have to demonstrate their competence by participation in proficiency testing.

Interlaboratory comparison (ILC) studies on the determination of PCDDs/PCDFs, dl-PCBs, indicator PCBs as well as selected BFRs in three natural food items have been performed annually since year 2000 by the Norwegian Institute of Public Health[1]. Up to 102 laboratories from 36 countries world-wide have reported results. This has provided a powerful tool for evaluating the quality and uncertainty in the analysis of dioxin-like compounds and other POPs in different food items such as meat, fish, eggs and dairy products[2]. Here we summarize the experiences during ten rounds of the ILC from 2003 on regarding the quality of POPs analyses in food world-wide.

Materials and methods

Study design

The analytes to be determined comprised 17 2,3,7,8-substituted PCDDs/Fs, 12 dioxin-like PCBs, 6 indicator PCBs, 8 PBDEs and 3 HBCDs (alternatively sum HBCDs). Analyses should be performed using the laboratories own methods for sample preparation and instrumental analysis, their own standards and quantification procedures, and their own method for lipid determination.

Sample collection, preparation and distribution

In each round, the test materials consisted of three non-spiked natural foodstuffs. In some cases, a highly contaminated product was mixed with a background contaminated product to obtain a reasonable contamination level. Homogenization of solid foodstuffs was performed by repeatedly grinding and mixing the food items. Liquid samples were stirred if necessary under elevated temperatures. Sub samples of the homogenates were placed into carefully cleaned screw-cap bottles and stored at -20 °C until shipment.

In order to assure homogeneity of the solid samples, we adopted an approach using electrolytic conductivity measurements after addition of sodium chloride. To measure background conductivity, boiling water was added to 10 g of homogenate and the mixture ultrasonicated. After centrifugation and filtration, conductivity was measured. To about 10% of the total homogenate, salt was added to approximately double the natural

conductivity. This sub-sample was added to the total sample prior to further homogenization of the total sample. Homogeneity of the test material was demonstrated by comparing the conductivity in water extracts of 10 samples from the same bottle (within bottle variation) and in extracts of 10 different bottles (between bottles variation).

For all rounds, standard solutions of all analytes in n-nonane were provided by Cambridge Isotope Laboratories, Inc., Andover, MA and shipped together with the food samples.

Reporting and handling of data

In 2003, the participants were asked to determine PCDDs/Fs and dl-PCBs. The following year, 8 PBDEs and the HBCDs were included in the study, and finally the 6 indicator PCBs were added in 2005. For each analyte in each sample, the participants were requested to report a single value for the concentration on wet weight basis alternatively indicate non-detect and report the detection limits. In addition, the determined concentrations in the standard solutions as well as the lipid content were to be reported. Each participating laboratory was given a laboratory code number. Participants had access to their own code only, and laboratory codes were not revealed to third parties.

Statistical analysis

An important objective of the ILC studies is to establish a good estimate of the true content of POPs in the different food items. In this respect, identification of outliers was an obvious challenge as the data set was far from normal distribution and often falsely high concentrations or LODs were reported. After evaluation of different approaches, the following method was found suitable for the calculation of the consensus concentrations for each of the PCDD/F and dl-PCB congeners:

Congener-by-congener medians were calculated from the data reported by all laboratories using the detection limit for non-detected congeners. Next, values two times above the medians were defined as outliers and removed from the data set. The consensus values were defined as the median of the remaining data set which was close to normally distributed thereby allowing calculation of the consensus mean and relative standard deviation (RSD). For indicator PCBs, PBDEs and HBCD, non-detected congeners were removed from the data set prior to consensus calculation. Outliers were defined as those values above two times the median of all values and were removed from the data set. The consensus values were defined as the median of the remaining data for each congener. For the standard solutions, outliers were defined as those values outside ± 50% of the median of all reported values. The consensus of the lipid content was calculated as the mean after removal of values outside ± 2 SD.

Toxic equivalencies (TEQs) were calculated from the consensus values for PCDD/Fs and dl-PCBs using the toxic equivalency factors (TEFs) derived by WHO in 1998 and 2005. Z-scores of TEQs as well as the sum of indicator PCBs, PBDEs and total HBCDs were calculated for each laboratory according to the following equation

$$Z = (x-X)/\sigma$$

where x = reported value; X = consensus value; σ = target value for standard deviation. A σ-value of 0.02 was chosen, i.e. Z-scores of +1 and -1 reflect deviations of 20% from the consensus value.

Results and discussion

Participation

The number of laboratories and their international provenance has increased during the years (Table 1).

	2003	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
No. of laboratories	77	73	87	96	100	102	92	89	101	90
No. of countries	24	24	28	28	27	33	31	31	31	36

Table 1. Number of reporting laboratories and countries for the ILC Study on POPs in Food since 2004.

In all rounds between 89 and 98% of the laboratories returned results. Participants were located world-wide representing up to 36 different countries, e.g. in 2012, 54 participants were from the European Union, 4 from other European countries, 13 from North America, 17 from East Asia, and 2 from Oceania.

Quality of reported data

Consensus statistics

Each report provides the statistics for the calculation of the consensus value. Usually, the median and mean consensus TEQs are in good agreement indicating a non-skewed distribution of results and hence an efficient removal of outliers. In Table 2 results are given for the 12th Round of ILC in 2011 where samples consisted of a highly contaminated fish sample, medium contaminated cheese and low contaminated egg. Nevertheless, the RSDs of the consensus mean in TEQ (outliers removed) are very similar and with about 12% quite low, demonstrating the generally good quality of the analytical work of the majority of laboratories. However, the range of TEQs which the laboratories would have reported based on their own results is very wide. Further, when comparing the consensus TEQs with the mean of TEQs reported by the laboratories, the latter are higher for two of the samples. This emphasizes the importance for laboratories to improve their analytical quality and participate in ILC studies in order to test their performance in the determination of dioxin-like compounds in food.

12th Round 2011	Median Total TEQ (pg/g fw)	Mean Total TEQ (pg/g fw)	RSD of TEQs (%)	Range of reported TEQs (pg/g fw)	Mean reported TEQs (pg/g fw)
Baltic salmon	8.1	7.9	13	0.47-91	8.7
Mozzarella cheese	1.2	1.1	11	0.026-4.8	1.2
Whole egg	0.76	0.77	12	0.015-7.8	0.88

Table 2. Statistical data of the consensus of total TEQs compared to TEQs reported by the laboratories.

For the 6 indicator PCBs and 8 PBDEs satisfying RSDs around 25% were obtained for the most abundant congeners (outliers removed), however, minor components such as CB-28, CB-52, CB-101 and BDE-28, BDE-187 had generally much higher RSDs.

Z-scores for PCDDs/Fs and dl-PCBs

The z-score is an estimate of the error in results scaled in standard deviation units (σ) where σ represents "the amount of uncertainty in the result that is tolerable in relation to the purpose of analysis" [3]. A σ-value of 20% for the TEQ z-scores was chosen to satisfy the performance criteria of the European Commission for methods of sampling and analysis for the official control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs[4]. For the last 10 rounds of the ILC, on average 2/3 of the participants had z-scores of ±1 for the PCDD/F TEQs and 83 % had z-scores of ±2. Z-scores of ±2 may be regarded as complying with the fitness-for-purpose criterion[3]. While there was no clear time trend in the performance of laboratories, the percentage within z-score ± 1 was dependent on the TEQ level (Figure 1).

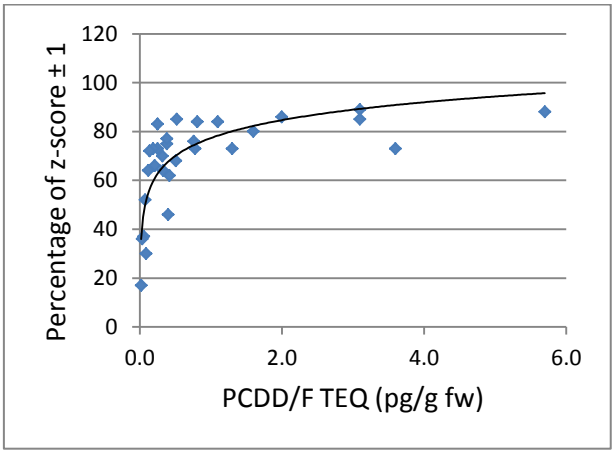

Figure 1. Fraction of laboratories with z-scores between ±1 as a function of PCDD/F TEQ.

It has been pointed out by Eppe et al.[5] that laboratories should not only assess their performance by a single result expressed in TEQ but in addition for all congeners that significantly contribute to the TEQ value. An accurate TEQ result is supposed to arise from accurate individual congener results and not by chance due to balanced effects between congeners. Further, the exact congener profile might be important for identifying sources by fingerprinting. However, in order to account for the dependency of the accuracy on the analyte concentration[6], a constant RSD of 20% is not appropriate. Eppe et al.[5] have therefore derived a "Dioxin function" which relates the expected precision of analytical result to the concentration of the analyte. Using this function to calculate σ-value for each individual congener's result, the percentage of z-scores between 1 is less dependent on the congener concentration.

Acknowledgements
We thank all the participating laboratories for good analytical work during the various rounds of the ILC on POPs in food and for their general interest in the overall objectives of these studies.

References
1. Norwegian Institute of Public Health, Interlaboratory Comparison on Persistent Organic Pollutants in Food, www.fhi.no/ilc, accessed 30.04.2013
2. Becher G, Haug LS, Thomsen C (2004) *Talanta* 63:1115-1122
3. Thompson M, Ellison SLR, Wood R (2006) *Pure Appl Chem* 78:145-196
4. European Commission (2012) *Official Journal* L 84:1-22
5. Eppe G, van Cleuwenbergen R, Haug LS, Boulanger B, Becher G, de Pauw E (2008); *Chemosphere* 71:379-387
6. Horwitz W, Kamps LR, Boyer KW (1980); *J Assoc Off Anal Chem* 63:1344-1354