Organohalogen Compounds, Volume 60, Pages 181-185 (2003)

ACCURACY AND COMPARABILITY OF ANALYTICAL DATA FOR PCDD/Fs AND PCBs IN FOODS

Martin Rose, Jim Startin

Defra Central Science Laboratory, Sand Hutton, **YORK** YO41 1LZ United Kingdom

Introduction

Wherever analytical data (or intake assessments etc. derived from them) on dioxins and dioxin-like chemicals are considered, the user needs to remain aware of various issues that may influence the accuracy and comparability of the results. Issues related to sampling, and to the accuracy of measured concentrations can have a significant effect on conclusions and inferences made from the data. Interpretation of reported data is complicated by the differing ways in which compounds that are not actually detected and measured are used. There are a disturbing number of statements in the literature about differences between foods or between locations that probably reflect simply differences in analytical performance and data assessment methodologies.

Analytical accuracy and precision

Standard protocols exist for calculating the measurement uncertainty associated with analytical measurements, such as the Eurachem Guide [1]. Such information including details of accuracy and precision is often absent from reports of data on PCDD/Fs and dioxin-like PCBs. The analysis of these substances in foods is particularly challenging because of the very low concentrations that are involved. Although bioassay tests [2,3] are becoming important, most data on dioxin-like compounds is obtained by a rather lengthy series of extraction, cleanup and concentration steps followed by HRGC-HRMS, involving instruments that require great skill in maintenance and operation. Several studies have shown that highly expert laboratories can achieve good agreement [4,5,6]. The number of laboratories engaged in these analyses has increased dramatically in the last few years but expertise is not necessarily acquired instantly. However, instrument sensitivity is improving and methods are more established; there is evidence that performance is generally improving. In the fourth round of WHO-coordinated interlaboratory quality assessment studies only three out of 11 laboratories participating met all of the quality criteria for the determination of PCDD/Fs in human milk [7], a less challenging task than is presented by cows' milk and some other foods. At the same time, in measurements of fat content less than half of the participants achieved deviations of less than 10% from the accepted value; many food data are expressed on a fat basis, the final result being calculated from measurements of the target compounds on a whole sample basis and separate measurement of the fat content.

More recent interlaboratory studies have been organised by the Norwegian Institute of Public Health. In 2000, the study included chicken, butter and fish, results being supplied by 37 laboratories from 15 countries [8]. In 2001, samples included beef, cod's liver and human milk, results being supplied by 55 laboratories from 24 countries [9]. Analytes included PCDD/Fs and dioxin-like PCBs, but not all laboratories determined both classes. In the second study it was shown that for levels of PCDD/Fs of 4 to 5 pg TEQ/g fat some 30 of the participating laboratories could determine the compounds with 'good accuracy' (± 20% of the consensus TEQ), but that for

samples with a lower content of PCDD/Fs (about 0.5 pg TEQ/g fat), considerably fewer laboratories (12) achieved this standard. It was noted that, because of the large number of non-detects, it was difficult to establish the true concentration and that the consensus value used as a target may actually have been an overestimate. For example, taking the consensus levels for the beef test material studied, in which the total TEQ (PCDD/Fs and PCBs) was 1.0 pg/g fat, 2378-TCDD, 12378-PeCDD and 23478-PeCDF together account for about 73% of the total TEQ for PCDD/Fs, and PCB 126 accounts for 73% of the total TEQ for PCBs. For 2378-TCDD, 29 of 52 reported values were non-detects. LODs were between 0.02 and 0.86 pg/g fat (an LOD of 10 was also reported but from a laboratory attempting the analysis with equipment that most analysts would regard as unsuitable). Positive results varied from 0.026 to 0.30 pg/g fat with a consensus of 0.1 pg/g fat. For 23478-PeCDF, for which the consensus value was 0.31 pg/g fat, there were only 7 non-detects, but results still ranged from 0.11 to 0.76 (excluding outlier of 27). For PCB 126, with a consensus value of 3.4 pg/g fat, results ranged from 0.3 to 8.1, although both these were identified as outliers and most results were within the range 1.0 - 5.8.

Limits of detection

It has been stated that in studies of food contaminants the analytical limit of detection (LOD) should be sufficiently low that further reduction of the LOD does not result in additional reduction in estimates of exposure [10]. For PCDD/Fs this criterion is met only in a very few laboratories, and even then not for every food under consideration. It is common for the concentrations of some or all of the individual dioxins and PCBs to be below the LOD. Inevitably, the sensitivity of analyses in different laboratories varies, and so does the method of assessing and reporting the LOD. Frequently, quantitative results are reported for any congener giving a GC-MS peak with a signal-to-noise ratio (S/N) of, say, 3 or more. If no peak is detected then an estimate is made of the concentration that would have produced this S/N. LODs so estimated vary between congeners and between different analyses. Some laboratories work, instead, to a 'limit of quantitation' (LOQ) which, strictly, should be defined as the lowest concentration at which a specified measurement precision is achieved, as demonstrated in method validation studies, but which is, instead, sometimes taken to be a multiple of the LOD. Alternatively, some data are reported after applying an arbitrary, but consistent, 'reporting limit'. Different laboratories analysing the same samples, or even interpreting the same raw data, may arrive at differing conclusions about which congeners are present at measurable concentrations, and at quite different LOD values for other congeners, even though their 'positive' results may be in good agreement. When data include non-detects then at least three different methods are in use for the calculation of total TEQ levels. These are the representation of non-detected congeners in subsequent calculations by: (a) a concentration of 0, (b) a concentration equal to the LOD or, (c) a concentration equal to ½ of the LOD. Even this may expand to five different methods if LODs and LOQs are differentiated.

The terms 'lower bound' and 'upper bound' are often used to refer to the first two of these methods, which correspond, respectively, to the minimum concentration *known* to be present and to a larger concentration, which *might* be present. The ND = ½ LOD calculation, for which the term 'median bound' is sometimes adopted, gives a result which lies midway between these upper and lower bounds. This often gives the appearance of better comparability, and may be preferable to the use of either lower or upper bound alone. However, if the upper bound and lower bound totals are far apart, as is frequently the case, then the mid-point is not necessarily any closer to the true total TEQ concentration than is either extreme. Use of any single convention in calculations

from data in which the LOD itself varies between analyses can be particularly misleading. Further complications may arise when results for specific congeners are averaged across a number of samples, with the total TEQ summation made using the average concentrations. In addition to use of all of the above substitution approaches, some workers average only detected values to generate a statistic, which is meaningless unless the frequency of detection is also taken into account.

All of these approaches have utility in some circumstances, if applied consistently. Upper bound estimates are favoured by the EU for regulatory monitoring, whilst the use of median bound is currently proposed for intake estimation and risk assessment purposes. In all cases, comparisons of total TEQ data representing different detection limits and calculation methods can be extremely misleading. There have undoubtedly been occasions when comparisons of total TEQs have been made and interpreted when nothing at all has been detected and measured, differences in the totals arising simply from differences in detection limit.

Sampling

The most accurate laboratory analysis can only give a result that represents the sample taken for analysis. How that sample is related to the broad food supply depends on both the quality and the intent of the sampling scheme. For use in estimation of intakes, sampling should be designed to take account of many factors, some of which are as follows: the proportion of each food that is imported, and variation in the countries of origin; seasonal variation, since supplies available to the consumer may be different throughout the year, imports may be more prevalent at different times, or food may have been stored for longer periods at out of season times; regional variation, as local food production may be effected by differences in climate or by local pollution sources or urbanisation, or by regional variation in available brands. Food prepared (‘take-away’ meals) or eaten outside the home (in restaurants, for example) also needs to be considered. Sampling to assess compliance with limits, whether statutory or ‘guidelines’, may not be appropriate for estimation of average intakes. For example, compliance monitoring may concentrate on domestic food production and exclude imports, or aim for broad geographical coverage without weighting by food production statistics. Any indication of localised contamination usually leads to more intensive sampling in the same location; small but significantly contaminated locations may therefore make a disproportionately large contribution to ‘average’ concentrations if these are based on all available data.

For most foodstuffs, achievement of a representative result inescapably necessitates coverage of a large number of samples. However, because of the cost and difficulty of analysis most of the older studies from the late 1980s and early 1990s, and some completed more recently, were based on rather limited numbers of samples, sometimes even single examples of a commodity or food product. One solution to this problem is to pool samples to form composites representing a specific category of foodstuff. This approach has been used in a number of national studies, for example in Finland, the Netherlands, New Zealand and the UK. The use of composite samples is a cost-effective way of obtaining robust measures of average concentrations but it does not furnish any information on the width and shape of the distribution of concentrations in the individual samples. It may also have the disadvantage of placing great reliance on single analyses.

Total diet study (TDS) schemes have often been used as the source of samples. In the UK TDS a total of 121 categories of food and drink are purchased fortnightly from 24 randomly selected locations representative of the UK as a whole. Samples are prepared and cooked as for

consumption and then combined into composite samples representing 20 defined food groups. The quantity and relative proportions of the foods that make up each composite are based on data from the National Food Survey and are updated annually [11]. Other total diet studies vary considerably in their geographical range, in the number of individual samples taken, and in the number and timing of samplings. Differences in food classification can also lead to lack of comparability between studies. In the UK TDS, the meat group is not segregated by animal species and includes beef, mutton and pork. In many studies some classes, such as meat products, fruit products, and cereals are differently defined and not comparable. Cereal, for example, may refer simply to grain and flour, or include cereal products encompassing breads, cakes and pastries prepared with animal fats, and sometimes various 'breakfast foods'.

Fat weight and whole weight reporting

PCDD/Fs and PCBs are lipid soluble and results for most food types containing over about 2% fat are reported on a fat weight basis. This gives more consistency for comparisons of samples such as milk, which show more variability with respect to dioxins on a whole weight basis than on a lipid basis. For some samples however, reporting on a fat-weight basis can lead to confusion. Fish can show seasonal variations in fat content, which can result in fat weight results giving an illusion of variation, even if the body burden with respect to dioxins remains constant. For low fat samples such as fruit and vegetables, the reported fat content can reflect organic co-extractives rather than true fat values. The amounts of these co-extractives is small and negligible for samples with higher fat content, but become significant when the fat content is very low. Results for these samples should only be considered on a whole-weight basis. Reporting of results on a fat-weight basis can also lead to different analytical approaches. Some laboratories isolate fat from samples (or ask for submission of isolated fat) before analysis starts. Internal standards are then added to the fat and the dioxins are determined on the fat sample. Other laboratories will add internal standards to the whole sample before isolation of fat, or will analyse the whole sample without fat isolation, and convert results to a fat-basis using the results of off-line determination of fat. Where there are inaccuracies in fat measurement or inefficient recoveries, these can lead to analytical differences.

Impact caused by the use of TEFs

It should be noted that TEF values are given only to the nearest one-half order of magnitude, and thus the range within which the 'true' TEF lies is right skewed, from ½ to 5 times its stated value. Treatment of data using probabilistic statistics techniques, which take into account uncertainty associated with assignment of TEFs, has shown that the TEQ can increase by 1.5 to 2.0 times compared to deterministic estimates using fixed values for the TEFs [12,13]. During the 1980s, a rather large number of different TEF schemes were used. International Toxic Equivalency Factors (I-TEFs) for PCDD/Fs were set in 1990 [14] and were adopted by almost all scientists and regulatory authorities. A more recent system of TEFs, set by the WHO in 1997 (WHO-TEFs) [15,16], has been accepted by most authorities and is coming into wider use, but much of the recent reporting of data for PCDD/Fs has continued to use the 1990 scheme. In the WHO-TEF scheme the TEF for 12378-PeCDD was doubled, from 0.5 in the I-TEF scheme to 1.0, while the TEFs for OCDD and OCDF were reduced by factors of 10, from 0.001 to 0.0001. The net effect of these changes for most food samples is an increase of around 15 to 20% in the calculated result for total TEQ level of PCDD/Fs. The WHO system also sets TEFs for those PCBs that bind to the Ah-receptor and elicit dioxin-like biochemical and toxic responses. In 1994, TEFs were set for

three non-ortho PCBs eight mono-ortho PCBs and for the di-ortho PCBs [17]. In the 1997 WHO scheme PCBs 170 and 180 were removed, PCB 81 added and the TEF for the non-ortho PCB 77 reduced by a factor of 5. For most food samples these changes make a negligible difference to the total TEQ attributable to PCBs. In many food samples the TEQ contribution made by PCBs may equal or, especially in fish, exceed that made by PCDD/Fs; it is therefore important to note whether or not PCBs are included in the total.

Conclusions

Great care needs to be taken when comparing sets of data that were generated in different laboratories or according to different quality criteria. Different quality and reporting criteria used in different laboratories, or even in the same laboratory at different times, can lead to false conclusions being drawn from analytical data.

References

1 Ellison, S., Rosslein, M. and Williams, A. (2000) *Quantifying Uncertainty in Analytical Measurement*. Eurachem/CITAC Guide

2 Murk, A.J.; Legler, J.; Denison, M.S.; Giesy, J.P.; van de Guchte, C.; Brouwer, A. *Fund. Appl. Toxicol.*, **33**, 149-160, 1996.

3 Bovee, T.F.H.; Hoogenboom, L.A.P.; Hamers, A.R.M.; Traag, W.A.; Zuidema, T.; Aarts, J.M.M.J.G.; Brouwer, A.; Kuiper, H.A. *Food Addit. Contam.*, **15**, 863-875, 1998.

4 de Jong, A.P.; Dross, A.; Fürst, P.; Lindstrom, G.; Päpke, O.; Startin, J.R. *Fresenius J. Anal. Chem.*, **345**, 72-77, 1993.

5 Malisch,R.; Schmid,P.; Frommberger,R.; Fürst, P. *Chemosphere*, **32**, 31-44, 1996.

6 Malisch, R.; Bruns-Weller, E.; Knoll, A.; Fürst P.; Mayer, R.; Wiesmuller, T. *Chemosphere*, **40**, 1033-1040, 2000.

7 World Health Organisation, Interlaboratory quality assessment of levels of PCBs, PCDDs and PCDFs in human milk and plasma: WHO Regional Office for Europe, Copenhagen, 2000.

8 Lindström, G.; Haug, L.S.; Nicolaysen, T., Intercalibration on dioxin in food - an international study, Report 2000:9, Statens institutt for folkehelse, Oslo, 2000. [http://www.folkehelsa.no/fag/miljoforu/edioxin.html]

9 Becher, G.; Lindström, G.; Nicolaysen, T.; Thomsen, C., Interlaboratory comparison on dioxin in food - second round of an international study, Report 2001:4, Statens institutt for folkehelse, Oslo, 2001. [http://www.folkehelsa.no/fag/miljoforu/edioxin.html]

10 Loftus, M.L.; Barraj, L.M.; Tomerlin, J.R. *J. AOAC Int.*, **75**, 911-915, 1992.

11 Peattie, M.E. *Food Chem. Toxicol.*, **21**, 503-507, 1983.

12 Smith, G.; Hart, A,; Rose, M.; MacArthur, R.; Fernandes, A.; White, S.; Moore, D. *Organohalogen Compounds*, **53,** 215-218, 2001.

13 Smith, G.; Hart, A,; Rose, M.; MacArthur, R.; Fernandes, A.; White, S.; Moore, D., *Food Addit. Contam.* 19 (8) 770-778, 2002.

14 Kutz, F.W.; Barnes, D.G.; Bottimore, D.P.; Greim, H.; Bretthauer, E.W., *Chemosphere*, **20**, 751-757, 1990.

15 van den Berg, M.; Birnbaum, L.; Bosveld, A.T.C.; Brunström, B.; Cook, P;, Feeley, M.; Giesy, J.; Hanberg, A.; Hasegawa, R.; Kennedy, S.W.; Kubiak, T.; Larsen, J.C.; van Leeuwen, R.F.X.; Liem, A.K.D.; Nolt, C.; Peterson, R.E.; Poellinger, L.; Safe, S.; Schrenk, D.; Tillitt, D.; Tysklind, M.; Younes, M.; Waern, F.; Zacharewski, T. *Environmental Health Perspectives*, **106**, 775-792, 1998.

16 van den Berg, M.; Peterson, R.E.; Schrenk, D. *Food Addit. Contam.*, **14**, 347-358, 2000.

17 Ahlborg, U.G. ; Becking, G.C.; Birnbaum, L.S.; Brouwer, A.; Derks, H.J.G.M.; Feeley, M.; Golor, G.; Hanberg, A.; Larsen, J.C.; Liem, A.K.D.; Safe, S.H.; Schlatter, C.; Wurn, F.; Younes, M.; Yrjanheikki, E. *Chemosphere*, **28**, 1049-1067, 1994.