IVM Institute for Environmental Studies

Third interlaboratory study on perfluorinated compounds in environmental and human matrices

Stefan van Leeuwen (IVM Institute for Environmental Studies, VU University, Amsterdam, The Netherlands)

Marie-Pierre Strub (Ineris, Verneuil-en-Halatte, France)

Wim Cofino (QUASIMEME Laboratory Performance Studies, Wageningen UR, Wageningen, The Netherlands)

Gunilla Lindström (MTM Research Center, Örebro University, Örebro, Sweden)

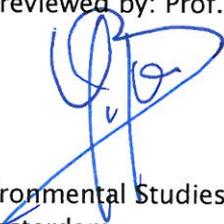Bert van Bavel (MTM Research Center, Örebro University, Örebro, Sweden)

This report is released by: Prof. Dr. Jacob de Boer
Head of Department Chemicals & Biology

INERIS

ÖREBRO UNIVERSITY

QUASIMEME

IIII ■ IVM Institute for
Environmental Studies

VU University Amsterdam

It was internally reviewed by: Prof. Dr. Jacob de Boer

IVM
Institute for Environmental Studies
VU University Amsterdam
De Boelelaan 1087
1081 HV AMSTERDAM
T   +31-20-598 9555
F   +31-20-598 9553
E   info@ivm.vu.nl

Contents

Summary

This international interlaboratory study on perfluorinated compounds (PFCs) in environmental samples was organized to assess the performance of laboratories world-wide on the analysis of PFCs human samples (human plasma samples A and B, 22 laboratories) and in environmental samples (water, fish and for the first time sludge, 40 laboratories). The participants used their in-house methods for analysis of the samples. Unknown solutions were also analysed by the participants to check their calibration procedures. The results were collected and statistically evaluated using the Cofino statistics. Z-scores were appointed to individual laboratory's results.

The samples A and B of the human study were identical and the results obtained corresponded nicely. Relative standard deviations (RSDs) of compounds with very well quantifiable levels in the plasma samples were 17 to 39% (except PFHpA), and consequently the majority of the laboratories (48-90%) obtained satisfactory Z-scores. The results for the unknown solutions were good and not unexpectedly better than the other samples.

The performance for the environmental samples was worse. A lower proportion of the laboratories obtained satisfactory Z-scores and RSD values were larger. This may partly be caused by the low levels of PFCs in the non-fortified samples. Other reasons are (i) some obvious outlier values were present and are presumably related to calculation errors and (ii) many laboratories used only a limited number of mass labelled internal standards, whereas it is recommended to use multiple. For the first time, a sewage sludge was included in the study. The variance for the results in this matrix was substantial, showing that more effort is needed to improve methods for sludge.

For PFOS specifically, sources contributing to the variation are (i) significant amounts of branched isomers present in the water, fish and blood, whereas calibration is often performed using only the linear isomer and (ii) some results reported might have been based on the salt rather than on the anion.

1   Introduction

Perfluorinated compounds (PFCs) are omnipresent in the environment (de Vijver *et al.* 2003; Smithwick *et al.* 2006; So *et al.* 2004). To study the distribution of these chemicals in the environment and to assess the environmental and human exposure, many laboratories have developed methods for analysis of PFCs in environmental matrices. For several years, the quality of data obtained was a major issue of concern (Martin *et al.* 2004). Problems identified in the quantification were the limited availability of high quality standards and mass labelled standards, severe matrix effects and interferences, the occurrence of branched isomers in industrial materials and blank problems due to contamination from labware and instrumentation. This was reflected in the poor results obtained in the 1st interlaboratory study (ILS) conducted in 2004/2005 on human and environmental matrices (van Leeuwen *et al.* 2006). Meanwhile, a large number of high quality standards became commercially available, as well as a wide range of mass labelled standards. A follow-up study on water and fish showed that accurate and precise analysis of PFCs in water and fish is feasible if several critical steps in the analysis are properly addressed, e.g. the use of high quality native standards and multiple mass labelled internal standards (van Leeuwen et al., 2009). Precise (i.e. low RSD values) and accurate results were obtained because all participants used the mass labelled internal standards that were provided in this study. A follow-up study on human serum in 2006 showed large improvements in the RSD between the participants. Also here the influence of the use of labelled standards for the quantification was clearly seen, in addition to more accurate quantification standards used (Lindström *et. al.* 2009). Further on more efficient sample extraction using among others solid phase extraction (SPE) improved the quality of the data.

This 3rd study was initiated to assess if the level of performance can be maintained. The number of participants increased compared to earlier studies. Forty laboratories participated in the environmental matrices study (of which 35 submitted data) and 22 in the human part of which all participants submitted data, but one laboratory submitted his data well after the set deadline. This data set is included in the report but has not been used for the statistical validation. The study focussed on the following PFCs: perfluorobutanoic acid (PFBA), perfluoropentanoic acid (PFPeA), perfluorohexanoic acid (PFHxA), perfluoroheptanoic acid (PFHpA), perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA), perfluoroundecanoic acid (PFUnA), perfluorododecanoic acid (PFDoA), perfluorotridecanoic acid (PFTrA), perfluorotetradecanoic acid (PFTeA), perfluorohexadecanoic acid (PFHxDA),perfluorooctadecanoic acid (PFODA), perfluorobutane sulfonate (PFBS), perfluorohexane sulfonate (PFHxS), perfluoroheptane sulfonate (PFHpS), perfluorooctane sulfonate (PFOS), perfluorodecane sulfonate (PFDS) and perfluorooctanesulfonamide (PFOSA).

In this study MTM research Centre (Örebro University) collaborated with the Institute for Environmental Studies (IVM) and QUASIMEME (www.QUASIMEME.org), INERIS (http://www.ineris.fr/) and the NORMAN network of reference laboratories for monitoring emerging environmental pollutants (www.norman-network.net).

2   Materials and methods

2.1   Material preparation

Water sample

A bulk water sample is taken from a local freshwater canal near Amsterdam in April 2009. After the suspended particulate matter (SPM) was allowed to settle, the bulk water was filtrated to remove small particles (<0.20 µm) and stored in a high density polyethylene (HDPE) tank while continuously mixing with a stainless steel mixing device. Individual HDPE bottles were filled with 500 ml sample prior to dispatch to the participants. All (bulk) sample handling and storage was performed at 4°C. The PFC concentrations in this sample may have been elevated due to a local accidental release of fire fighting foam in July 2008.

Fish sample

The fish sample was obtained from QUASIMEME Laboratory Performance Studies (LPS). The selected fish was a pike perch (*Stizostedion lucioperca*), sampled in Lake IJssel, The Netherlands. The muscle tissue is collected by filleting the fish. The bulk material was ground and thoroughly homogenised after addition of 0.02% butylhydroxytoluene (BHT) as an antioxidant. Individual jars were filled with approx. 65 grams of muscle homogenate. The jars were sterilized at 3 bar, 120°C, which allowed storage at room temperature and convenient transportation. Details on the preparation process of similar materials can be found elsewhere (de Boer 1997).

Sludge sample

The sludge sample is provided by WEPAL (www.wepal.nl). The sludge originates from the Netherlands. The bulk sludge material was dried at 40°C and milled to pass a 0.5 mm sieve. The bulk was homogenised and individual bottles were filled with sample material.

Plasma sample

Both human plasma samples were prepared from plasma pools collected from 100 man and women in the US. All material was homogenized and 1ml packed in pre-cleaned glass bottles at NIST. The material will be made available as SRM by NIST after the study and the statistical evaluation. The samples should preferably be stored frozen until analysis. The PFC concentrations should reflect the levels in the normal (US) population and some of the PFCs might not be present or present at very low concentrations.

Unknown solutions IVM PFC-MXA and PFC-MX3

Both unknown solutions contained 13 perfluorinated carboxylic acids (PFCAs) and 4 perfluorinated sulfonic acids (PFSAs) in methanol. The concentrations of the IVM PFC-MXA were in the range of 20-60 ng/mL for all compounds and the concentrations of PFC-MX3 were in the range of 80-280 ng/ml. It should be noted that the certified PFSA concentrations are based on the sodium or potassium salts. Details on the concentrations of the individual PFCs can be found in Table 1 and 2 and in Appendix 2.

Mixture of mass labelled PFCs MPFAC-MXA-100

This ampoule (with approx. 0.6 ml) contained 6 mass labelled PFCAs and 2 PFSAs in methanol. These standards were provided to the participants for optional use as internal standards. Participants were free to use these standards or own standards or spiking protocols to achieve the best possible results.

2.2   Methods used by participants

For the different sample matrices the participant have used different methods fine tuned to achieve the best results. In Appendix 5 the methods reported by each individual participant are given. In Figure 1 and 2 the methods used are summarised. For the liquid samples (plasma and water) the preferred method is solid phase extraction (SPE). This combined extraction and clean up method is predominantly used for nearly 80% of the water samples. A standard method for water samples using SPE is available (ISO 25101:2009) and this method or similar methodology has been used. In addition direct injection after diluting the sample was used by some of the participants beside the traditional liquid/liquid extraction.

The methods used for the plasma samples are more diverse: ion pair, SPE, sample dilution, carbon dispersion and on-line injection have been used. When comparing the methods for plasma samples with earlier Fluoros QA/QC studies, the increase in use of on-line methods is striking.
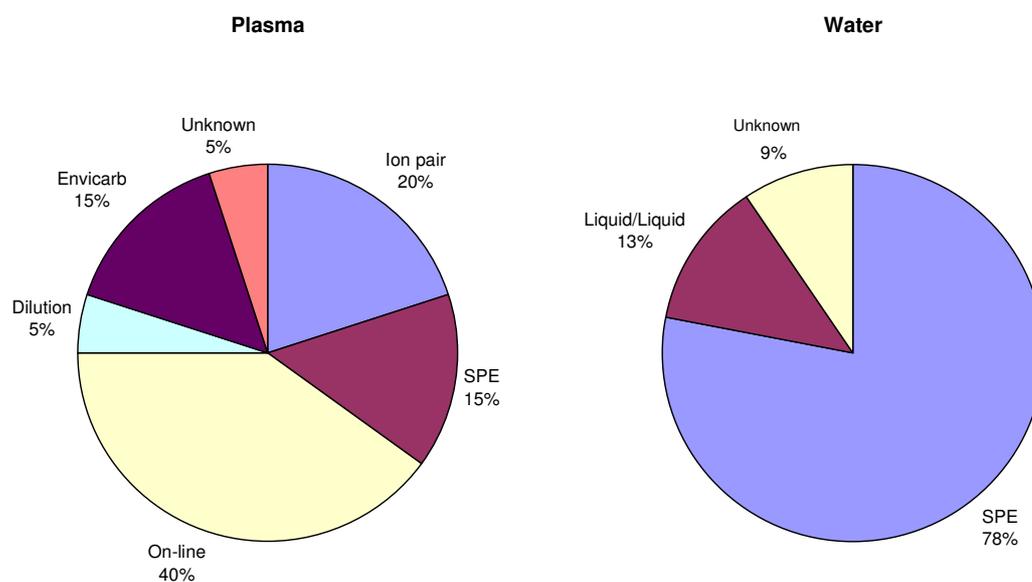

Figure 1   *Methods used for plasma and water samples, for details see Appendix 5*

For the fish samples most of the participants used liquid extraction with acetonitrile or methanol, only a limited number of laboratories have used ion pair or direct SPE extraction. After extraction the sample is injected directly by 35% of the participants, but the majority uses SPE or activated carbon (Envicarb) to further clean up the extract and thus reduce interferences during the final analysis.
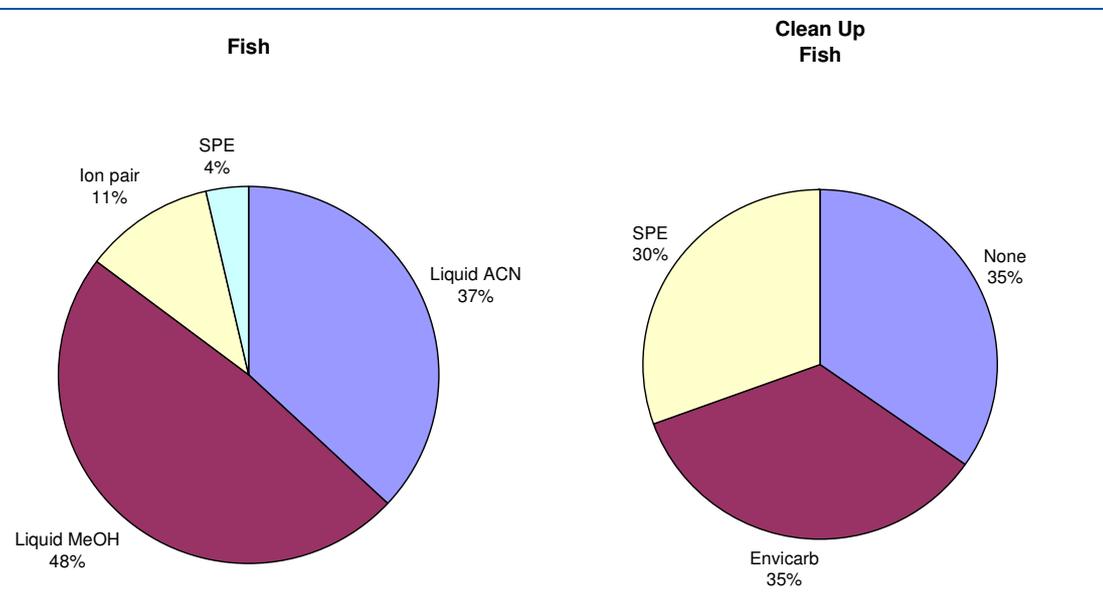

Figure 2   Method and clean up used for the fish sample, for details see Appendix 5

For the sewage sludge sample, most participants (approx 80%) used a method very similar to (their) fish extraction and clean-up methods (i.e. extraction with methanol or acetonitrile, clean-up with Envicarb or anion-exchange SPE). A couple of labs included an acid or base (e.g. NaOH) step in their extraction procedure specifically for the sludge matrix.

The final detection of the target compounds was liquid chromatography coupled to mass spectrometry. The majority of the laboratories used LC triple quad MS (LC/MS/MS) but a few laboratories used ion trap (LC/ITMS) or time of flight (LC/QTOF). Nearly all participants used mass labelled internal standards but not for all target compounds. Several laboratories used the provided labelled standard which contained six mass labelled compounds or at least mass labelled PFOS and PFOA. Surprisingly 3-4 laboratories, depending on the sample matrix, did not use an internal standard at all. About half of the participants corrected for recovery, although it is somewhat unclear if this was done using the labelled internal standards or spiking experiments. Most of the results were not corrected for blank levels, no further information on the blank correction was given. A limited number of laboratories used matrix matched standard solutions to reduce the effect of the matrix, the majority of the labs however ran the standards in methanol, acetonitrile or a mixture of both solvents and water.

2.3   Data Assessment

The data assessment was carried out according to the principles employed in the data assessment of the QUASIMEME LPS. All data received from the participants were entered into a database and assessed using a standard procedure to allow direct comparison between participants. The approach to the assessment is based on a

standard, ISO 13528 (2005), the IUPAC International Harmonised Protocol for Proficiency Testing (Advanced Draft) by Thompson et al. (2006). Additions or differences in the assessment from these standards are given or referred to in this report. However, the assigned value and the laboratory assessment using Z-scores are based on the Cofino Model (Cofino et al., 2000).

Comparison between the robust statistics method for calculation of a mean and the Cofino model continues to be made, and where there are any significant discrepancies between the two methods then further investigative analysis was undertaken. The Cofino model is generally able to separate the effects of the analytical method on the results and provide a more reliable estimate of the measurement relating to the method. The standard, ISO 13528, includes statistics for proficiency testing schemes, and uses robust statistics as a basis for the assessment. However, it is generally acknowledged that robust statistics cannot cope with more than 10% extreme values, particularly with a skewed distribution. The Cofino model is able to routinely cope with these types of distribution and provide the best estimate of the consensus value, which may be used as the assigned value.

The Cofino model has been developed for the routine QUASIMEME assessments. The Cofino model uses a Normal Distribution Assumption (NDA). The assigned value is based on the Cofino NDA model without any trimming of the data. This approach includes all data in the evaluation and no subjective truncation or trimming is made. This model has been further developed to include Left Censored Values (LCV)[1]. The development of these models has been fully documented and published (Cofino et al, 2000; Cofino et al., 2005; Wells et al., 2004). An overview of the assessment with explanation and examples is given in the Assessment Rules for the Evaluation of the QUASIMEME LP Studies Data (Wells and Scurfield, 2004).

The details of the Cofino Model were provided elsewhere (Wells et al., 2004; Wells and Scurfield, 2004) but in summary the approach is as follows:

- All data included in the assessment
- No data trimmed or down weighted
- Assigned values (AV) based on Cofino NDA model
- All LCV are also included, provided certain criteria are met

2.3.1   Plots

The performance of the laboratories in this study is illustrated in the Z-score histograms in Figure 3. Where the assigned value for a determinand is indicative, the values are plotted as their original reported concentrations. The rules for confirming whether the consensus value should be an assigned value or an indicative value are given in the Assessment Rules for the Evaluation of the QUASIMEME LP Studies Data (Wells and Scurfield, 2004) with appropriate examples.

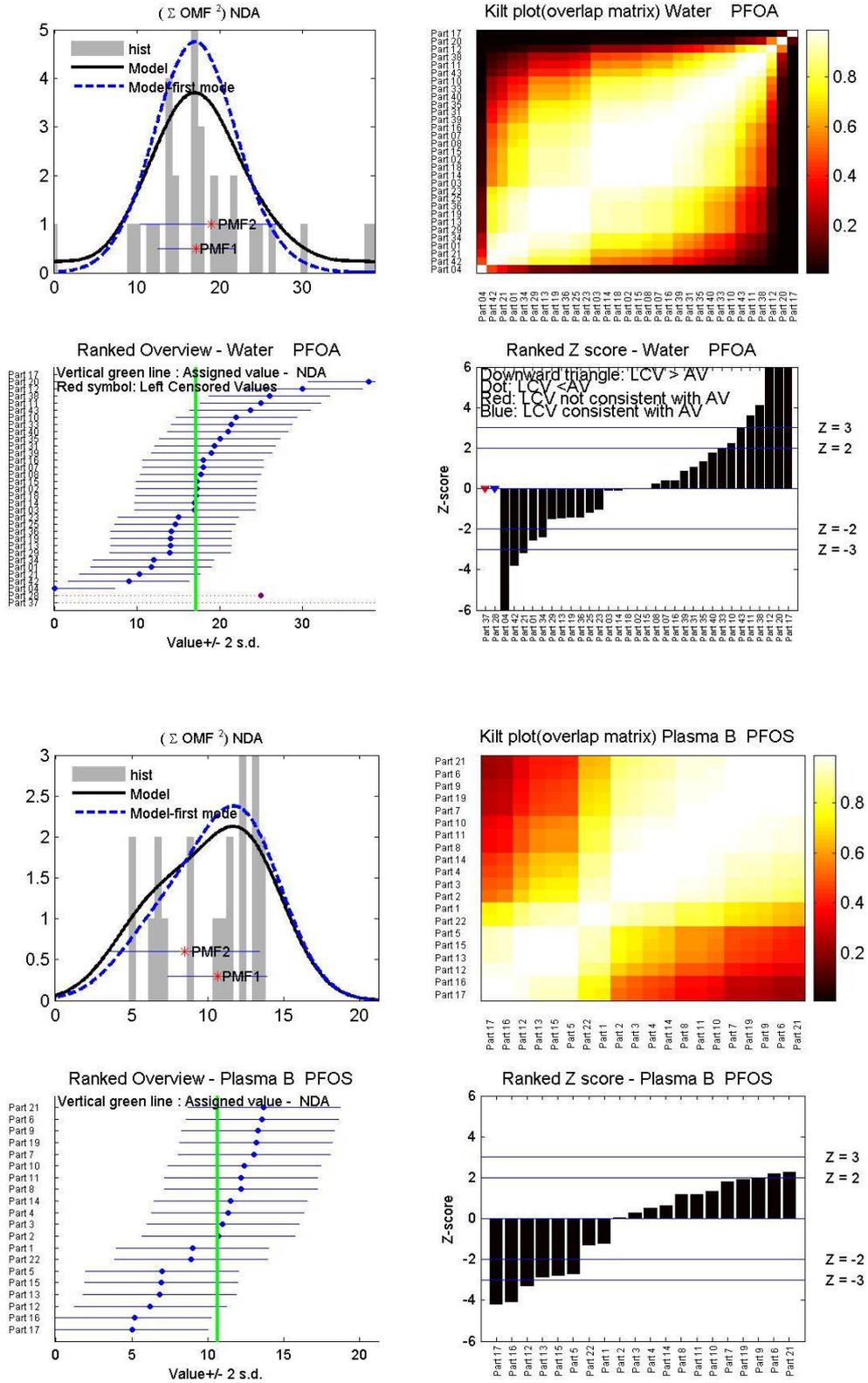[1]   *Left Censored Values* is the correct nomenclature for "less than" values

Figure 3   Examples of the graphical output of the Cofino Model statistics for PFOA in water (top) and PFOS in the plasma B sample (bottom)

Normally, four plots are given for each determinand (Figure 3). The upper left plot provides an impression of the probability density function (PDF) for all data (black) and for the first mode (blue dotted) (PMF1) of the data. Superimposed on these PDFs is a histogram of the individual measurements, given in grey. This plot shows the distribution of the data as a whole, and of the data in the main mode (PMF1) on which the assigned value is based.

The "Kilt Plot" (Overlap Matrix) (upper right plot) provides an overview of the degree of overlap of each pair of data. It gives a clear indication of the degree of homogeneity of the data. As a key, the white areas indicate maximum overlap of the PDFs and, therefore, highest agreement (an overlap of one implies that the two laboratories of the pair report exactly the same results), while the black area show the pairs in poor agreement.

The lower left plot is a ranked overview of all data with an error bar of ± 2 SD. The numerical values are given in blue and the left censored values are given in red.

The ranked Z-score plot (lower right) is based on the mean of the data, which is normally also the assigned value. However, if there is any adjustment required to the assigned value as a result of the assessment, e.g. use of the nominal concentration or a trimmed value, then the final Z-score given in the Z-score histograms will reflect these changes. In this study, no such adjustments are made and therefore, the Z-score plot (lower right) is the definite plot for obtaining the individual lab Z-scores.

2.3.2   The Assigned Values and indicative values

The Assigned Value (AV) is obtained from the main mode of the data using the Cofino Model (bleu dotted line in upper left panel in Figure 3), and is cantered around the highest density of values. Unless otherwise stated, the assigned value is based on this consensus value of *all* data. Although *all* data are included in the assessment, those values that lie some distance from AV contribute less to the mean than values which occur at or near the mean.

In some instances it is not possible to set an AV, and an indicative value is given. No assessment of laboratory performance is given where an indicative value is set. An overview of the assessment, with explanation, decision flowcharts and examples, is given in the paper *Assessment Rules for the evaluation of the QUASIMEME Laboratory Performance Studies Data*, available on the QUASIMEME website, www.quasimeme.org. A summary of the categories is given below:

Category 1

For data with the number of numerical observations ≥ 7

An assigned value is based on the mean when ≥ 33% of values have a Z-score of $|Z| < 2$. Where < 33% of the data has $|Z| < 2$ the value is indicative. i.e. at least 33% must be in good agreement.

Category 2

For data with the number of numerical observations > 3 and < 7

An assigned value is based on the mean when ≥ 70% of values have a Z-score of $|Z| < 3$ and a minimum of 4 observations have $|Z| < 2$. Otherwise the value is indicative. i.e. for small datasets, n > 3 and n < 7, there needs to be very good agreement and a maximum of one extreme value before an assigned value can be given.

Category 3

For data with the number of numerical observations < 4

No assigned value is given. Normally the median value is given as an indicative value.

Category 4

For data with the high Total Error% >100% in combination with bad performance, no assigned value is given.

2.3.3   The Z-score Assessment

A Z-score (Thompson and Wood, 1993) is calculated for each participant's data for each matrix / determinand combination which is given an assigned value. The Z-score is calculated as follows:

$$\text{Z-score} = \frac{\text{Mean from Laboratory - Assigned Value}}{\text{Total Error}}$$

It is emphasized that in many interlaboratory studies the between-laboratory standard deviation obtained from the statistical evaluation of the study is used as 'total error' in the formula above. In the QUASIMEME data assessment, the total error is estimated independently taking the needs of present-day international monitoring programs as starting point. For each determinand in a particular matrix, a proportional error (PE) and a constant error (CE) have been defined. The total error depends on the magnitudes of these errors and on the assigned value:

$$\text{Total Error} = \frac{\text{Assigned Value} \times \text{Proportional Error (\%)}}{100} + 0.5 \times \text{Constant Error}$$

The values for the PE and CE are set by the QUASIMEME Scientific Assessment Group and are monitored annually. The values are based on the following criteria:

• Consistency of the required standard of performance to enable participating laboratories to monitor their assessment over time.

• Achievable targets in relation to the current state of the art and the level of performance needed for national and international monitoring programmes.

The assessment is based on ISO 17043 as Z-scores. The QUASIMEME model is designed to provide a consistent interpretation over the whole range of concentration of analytes provided, including an assessment where Left Censored Values (LCVs) are reported.

The proportional error is set at 12.5% for all matrices. This applies to all determinands. The constant error (CE) has been set for each determinand or determinand group (e.g. chlorinated biphenyls). This value was initially set to reflect the limit of determination, but is at present more closely related to the overall laboratory performance. The magnitude of the CE is set to provide a constant assessment in terms of Z-score regardless of concentration. The CE is set at 0.025. Therefore at low concentrations the level of accuracy required to obtain a satisfactory Z-score is less stringent than at a high concentrations.

Following usual practices e.g. ISO 17043, the Z-scores can be interpreted as follows for laboratories which take part in QUASIMEME to assure the quality of their data for use in international marine monitoring programmes:

$|Z| < 2$ Satisfactory performance

$2 < |Z| < 3$ Questionable performance

$|Z| > 3$ Unsatisfactory performance

The following schematic presentation illustrates the interpretation of the Z-scores:
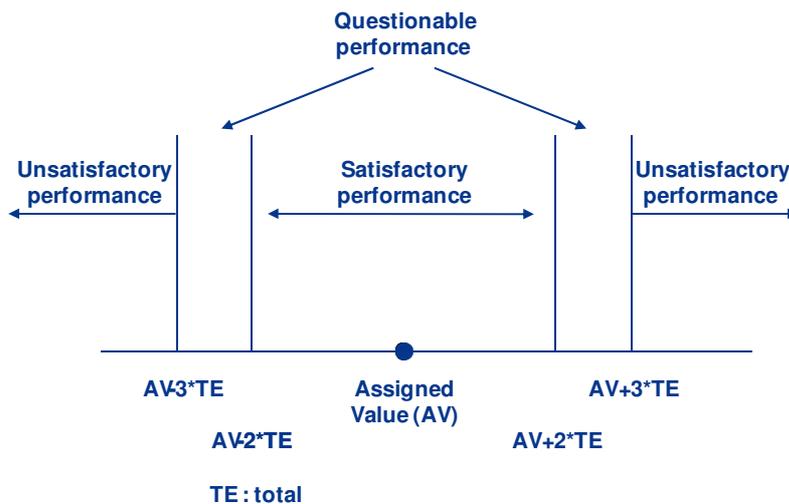
$|Z| > 6$ frequently points to gross errors (mistakes with units during reporting, calculation or dilution errors, and so on).

It is not possible to calculate a Z-score for left censored values (LCV's) as LCVs represent a cut-off value rather than continuous data. However, Quasimeme provides a simple quality criterion:

LCV/2 < (concentration corresponding to $|z|=3$): LCV consistent with assigned value

LCV/2 > (concentration corresponding to $|z|=3$): LCV inconsistent with assigned value, i.e. LCV reported by laboratory much higher than numerical values reported by other laboratories.

Z-score key: S – Satisfactory

Q – Questionable

U – Unsatisfactory

LCV key: C – Consistent

I – Inconsistent

No data: B - Blanc

3   Results

The submitted results have been evaluated statistically and whenever the data met the requirements (as mentioned in chapter 2), an assigned value was established. Z-scores were calculated based on the assigned value. Summary statistics are presented in Table 1 and 2. A summary of the assigned values and the percentage of satisfactory to unsatisfactory Z-scores is presented in Table 3 and 4. Whenever numerical less than values (left censored values, LCV) were submitted, it is mentioned whether these LCVs are consistent with the assigned value. The submitted data is presented in Appendix 2. Tables with individual Z-scores are presented in Appendix 3 and Z-score plots in Appendix 4.

Table 1a.   Summary of results of PFCs in the environmental samples (unknown solution, results in ng/ml)

Unknown solution)	Design Value	Assigned value	Average	Median	Min.	Max.	SD	%RSD	n*
PFBA	60.0	59.6	58.1	58.9	22.9	95.3	14.9	26	19
PFPeA	30.0	29.4	29.4	28.6	12.9	46.7	7.7	26	23
PFHxA	40.0	39.1	45.7	39.1	23.0	233	36.8	81	29
PFHpA	25.0	24.3	25.0	25.0	15.5	39.8	4.9	20	26
PFOA	70.0	72.4	81.8	73.0	48.1	411	58.7	72	35
PFNA	40.0	40.8	43.3	40.7	25.4	106	13.7	32	28
PFDA	30.0	30.0	38.3	30.0	6.7	258	42.5	111	30
PFUdA	20.0	20.2	30.4	20.3	7.3	245	45.2	149	25
PFDoA	20.0	20.0	30.9	20.2	6.3	285	54.4	176	24
PFTrDA	20.0	18.5	31.2	19.4	6.4	175	48.0	154	11
PFTeDA	20.0	18.3	20.4	20.2	9.7	42.3	9.5	46	10
PFHxDA	20.0	NA	23.0	23.0	NA	NA	NA	NA	1
PFODA	20.0	NA	22.0	22.0	NA	NA	NA	NA	1
L-PFBS	35.4 (40.0)**	36.6	37.4	36.0	17.7	77.0	11.5	31	26
L-PFHxS	28.4 (30.0) **	28.5	29.1	28.6	11.6	65.0	9.7	33	27
L-PFOS	55.6 (60.0) **	57.8	61.9	56.1	25.8	178	24.0	39	35
L-PFDS	19.3 (20.0) **	18.5	21.4	18.7	3.0	43.2	10.1	47	12

*   n: number of submitted datasets
**   Concentration of the anion and concentration of the salt between brackets.

Table 1b.   Summary of results of PFCs in the environmental study samples (fish, results in ng/g ww)

Fish	Assigned value	Average	Median	Min.	Max.	SD	%RSD	n*
PFBA	NA	0.16	0.16	NA	NA	NA	NA	1
PFPeA	NA	0.24	0.24	NA	NA	NA	NA	1
PFHxA	NA	2.07	2.15	0.09	3.90	1.56	75	4
PFHpA	NA	1.28	1.00	0.80	2.05	0.67	52	3
PFOA	NA	3.80	0.39	0.09	33.0	10.3	270	10
PFNA	0.52	1.23	0.60	0.23	6.75	1.63	132	19
PFDA	2.62	4.04	2.69	0.66	27.0	5.11	127	24
PFUdA	1.43	1.73	1.40	0.38	4.70	1.12	65	19
PFDoA	0.27	1.05	0.30	0.20	5.30	1.57	149	13
PFTrDA	NA	0.73	0.39	0.10	2.16	0.83	114	5
PFTeDA	NA	0.08	0.08	NA	NA	NA	NA	1
PFHxDA	NA	NA	NA	NA	NA	NA	NA	NA
PFODA	NA	NA	NA	NA	NA	NA	NA	NA
L-PFBS	NA	0.10	0.10	0.00	0.00	0.14	141	2
L-PFHxS	NA	0.89	0.11	0.00	5.10	1.87	210	7
L-PFOS	65.4	61.6	67.0	2.48	110	24.8	40	27
L-PFDS	NA	1.93	0.29	0.14	5.35	2.96	154	3
PFOSA	1.44	1.80	1.60	0.88	3.60	0.99	55	8

*   n: number of submitted datasets

Table 1c.   Summary of results of PFCs in the environmental study samples (water, results in ng/L)

Water	Assigned value	Average	Median	Min.	Max.	SD	%RSD	n*
PFBA	12.2	12.48	12.06	6.70	20.05	12.48	37	8
PFPeA	3.76	5.06	4.50	2.62	10.40	5.06	51	9
PFHxA	9.00	8.80	9.00	0.00	16.50	8.80	36	22
PFHpA	3.48	3.72	3.66	1.99	7.30	3.72	36	17
PFOA	17.2	18.92	17.20	0.00	53.00	18.92	49	31
PFNA	0.93	0.90	0.94	0.40	1.47	0.90	37	14
PFDA	NA	6.18	0.95	0.33	28.00	6.18	160	11
PFUdA	NA	11.20	11.20	0.40	22.00	11.20	136	2
PFDoA	NA	1.33	0.40	0.00	3.60	1.33	148	3
PFTrDA	NA	NA	NA	NA	NA	NA	NA	0
PFTeDA	NA	NA	NA	NA	NA	NA	NA	0
PFHxDA	NA	NA	NA	NA	NA	NA	NA	0
PFODA	NA	NA	NA	NA	NA	NA	NA	0
L-PFBS	8.56	10.10	8.50	4.40	31.00	10.10	61	17
L-PFHxS	24.9	23.45	24.85	0.00	43.00	23.45	35	26
L-PFOS	75.0	78.83	73.40	0.00	180	78.83	46	35
L-PFDS	NA	NA	NA	NA	NA	NA	NA	0
PFOSA	NA	1.87	1.87	0.24	3.50	1.87	123	2

*   n: number of submitted datasets

Table 1d.   Summary of results of PFCs in the environmental study samples (sludge, results in ng/g)

Sludge	Assigned value	Average	Median	Min	Max	SD	%RSD	n*
PFBA	NA	16.84	22.90	0.62	27.00	14.20	84	3
PFPeA	NA	3.06	3.12	0.77	5.00	1.59	52	7
PFHxA	0.84	7.33	0.99	0.37	61.31	19.02	259	10
PFHpA	NA	1.37	0.65	0.21	8.00	2.14	156	12
PFOA	10.7	11.88	10.60	2.00	38.00	6.93	58	25
PFNA	0.39	1.12	0.44	0.22	4.50	1.57	139	12
PFDA	2.58	3.12	2.61	0.80	11.30	2.28	73	18
PFUdA	NA	1.37	1.00	0.19	4.17	1.15	84	13
PFDoA	1.89	1.99	1.90	0.45	3.90	0.98	49	11
PFTrDA	NA	0.20	0.20	NA	NA	NA	NA	1
PFTeDA	NA	0.20	0.20	NA	NA	NA	NA	1
PFHxDA	NA	NA	NA	NA	NA	NA	NA	0
PFODA	NA	NA	NA	NA	NA	NA	NA	0
L-PFBS	NA	2.28	1.00	0.00	8.50	2.93	128	7
L-PFHxS	NA	1.67	1.40	0.24	4.46	1.28	77	12
L-PFOS	89.3	87.07	87.00	1.14	152	40.68	47	27
L-PFDS	7.03	8.71	7.76	3.30	18.63	5.41	62	8
PFOSA	NA	3.83	4.15	0.51	6.20	1.86	49	6

*   n: number of submitted datasets

Table 2a.   Summary of results of PFCs in the human study samples (unknown solution, results in ng/ml)

Unknown solution	Design Value	Assigned value	Average	Median	Min	Max	SD	% RSD	n*
PFBA	240	222.	206	222	122	278	42	21%	16
PFPeA	120	116	108	116	54	159	31	29%	15
PFHxA	160	144	143	147	79	213	35	24%	18
PFHpA	100	99.5	98	99	51	145	23	24%	20
PFOA	280	252	267	256	207	546	73	27%	21
PFNA	160	153	152	155	98	202	30	20%	20
PFDA	120	118	110	117	27	150	28	25%	20
PFUdA	80	77.7	78	79	46	123	18	23%	19
PFDoA	80	71.2	73	74	23	142	25	35%	19
PFTrDA	80	79.2	94	83	70	179	34	36%	11
PFTeDA	80	84.0	101	86	71	186	36	35%	12
PFHxDA	80	NA	127	148	77	156	43	34%	3
PFODA	80	NA	93	95	77	107	15	17%	3
L-PFBS	142 (160)*	138	132	141	59	190	35	26%	15
L-PFHxS	114 (120)*	107	103	110	57	150	19	19%	20
L-PFOS	223 (240)*	214	218	219	120	419	64	29%	21
L-PFDS	77 (80)*	83.9	86	90	18	152	39	46%	11

*   n: number of submitted datasets
**   Concentration of the anion and concentration of the salt between brackets.

Table 2b.   Summary of results of PFCs in the human study samples (plasma A, results in ng/ml)

Plasma A	Design Value	Assigned value	Average	Median	Min	Max	SD	%RSD	n*
PFBA	NA	NA	0.20	0.20	NA	NA	NA	NA	1
PFPeA	NA	NA	1.35	1.35	0.09	2.60	1.77	132%	2
PFHxA	NA	NA	0.14	0.07	0.03	0.33	0.16	114%	3
PFHpA	NA	0.18	0.32	0.19	0.14	1.70	0.41	127%	14
PFOA	NA	2.87	2.98	2.94	2.00	4.87	0.63	21%	22
PFNA	NA	0.82	0.81	0.84	0.44	1.22	0.23	29%	19
PFDA	NA	0.22	0.23	0.22	0.11	0.39	0.08	35%	15
PFUdA	NA	0.13	0.70	0.15	0.05	5.00	1.61	229%	9
PFDoA	NA	NA	1.85	0.23	0.03	6.90	3.37	182%	4
PFTrDA	NA	NA	0.05	0.05	0.04	0.07	0.024	45%	2
PFTeDA	NA	NA	0.13	0.13	NA	NA	NA	NA	1
PFHxDA	NA	NA	NA	NA	NA	NA	NA	NA	0
PFODA	NA	NA	NA	NA	NA	NA	NA	NA	0
L-PFBS	NA	NA	0.03	0.03	0.01	0.06	0.036	105%	2
L-PFHxS	NA	3.33	3.28	3.38	1.41	4.90	0.80	25%	21
L-PFOS	NA	10.6	9.95	11.31	4.55	14.00	3.29	33%	22
L-PFDS	NA	NA	0.002	0.002	NA	NA	NA	NA	1

*   n: number of submitted datasets.

Table 2c.   Summary of results of PFCs in the human study samples (plasma B, results in ng/ml)

Plasma B	Design Value	Assigned value	Average	Median	Min	Max	SD	%RSD	n*
PFBA	NA	NA	0.20	0.20	NA	NA	NA	NA	1
PFPeA	NA	NA	1.04	0.13	0.08	2.90	1.61	156%	3
PFHxA	NA	NA	0.12	0.07	0.04	0.26	0.12	97%	3
PFHpA	NA	0.17	0.26	0.18	0.10	1.10	0.25	98%	14
PFOA	NA	2.94	2.97	3.05	1.97	4.43	0.58	19%	21
PFNA	NA	0.83	0.83	0.86	0.52	1.20	0.19	23%	20
PFDA	NA	0.24	0.23	0.25	0.11	0.35	0.07	29%	15
PFUdA	NA	0.14	0.65	0.15	0.05	5.20	1.60	247%	10
PFDoA	NA	NA	1.81	0.12	0.03	7.00	3.46	191%	4
PFTrDA	NA	NA	0.03	0.03	0.004	0.05	0.033	120%	2
PFTeDA	NA	NA	0.01	0.01	NA	NA	NA	NA	1
PFHxDA	NA	NA	NA	NA	NA	NA	NA	NA	0
PFODA	NA	NA	NA	NA	NA	NA	NA	NA	0
L-PFBS	NA	NA	0.02	0.02	NA	NA	NA	NA	1
L-PFHxS	NA	3.27	3.30	3.15	1.39	5.60	0.89	27%	21
L-PFOS	NA	10.6	10.16	11.17	5.01	13.70	3.00	29%	21
L-PFDS	NA	NA	0.003	0.003	NA	NA	NA	NA	1

*   n: number of submitted datasets.

Table 3a.   Summary of laboratory performance for PFCs in environmental matrices (fish, results in ng/g ww)

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satis-factory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% consis-tent LCV	% inconsis-tent LCV
PFBA	NA		22						
PFPeA	NA		28						
PFHxA	NA		39						
PFHpA	NA		44						
PFOA	NA		61						
PFNA	0.5	14.89	61	36	18	9	23	5	9
PFDA	2.6	12.98	67	50	4	21	25		
PFUdA	1.4	13.37	56	55	15	15	10	5	
PFDoA	0.3	17.22	50	50			22		28
PFTrDA	NA		14						
PFTeDA	NA		14						
PFHxDA	NA								
PFODA	NA								
PFBS	NA		36						
PFHxS	NA		44						
PFOS	65.4	12.52	75	59	7	30	4		
PFDS	NA		22						
PFOSA	1.4	13.37	22	50	25		25		

Table 3b.   Summary of laboratory performance for PFCs in environmental matrices (water, results in ng/L

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satis-factory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% consis-tent LCV	% incon-sistent LCV
PFBA	12.2	12.60	31	36	9	27		18	
PFPeA	3.8	12.83	42	27	13	7	13	27	7
PFHxA	9.0	12.64	72	50	23	4	8	8	4
PFHpA	3.5	12.86	64	43	9	17	4	13	9
PFOA	17.2	12.57	94	56	9	18	9	3	3
PFNA	0.9	13.85	61	36	9	18		9	27
PFDA	NA		67						
PFUdA	NA		50						
PFDoA	NA		50						
PFTrDA	NA		14						
PFTeDA	NA		19						
PFHxDA	NA								
PFODA	NA								
PFBS	8.6	12.65	58	38	5	33	5	10	10
PFHxS	24.9	12.55	75	70	7	15	4		4
PFOS	75.0	12.52	100	50	22	11	14	3	
PFDS	NA		19						
PFOSA	NA		19						

Table 3c.   Summary of laboratory performance for PFCs in environmental matrices (sludge, results in ng/g dw)

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satis-factory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% consis-tent LCV	% inconsis-tent LCV
PFBA	NA		19						
PFPeA	NA		36						
PFHxA	0.8	13.99	47	24	6	18	12	18	18
PFHpA	NA		50						
PFOA	10.7	12.62	75	59	7	15	11		4
PFNA	0.4	15.70	47	41	12		18	12	12
PFDA	2.6	12.98	58	57	10	10	10	10	
PFUdA	NA		44						
PFDoA	1.9	13.16	42	40	7	20	7	13	7
PFTrDA	NA		11						
PFTeDA	NA		14						
PFHxDA	NA		3						
PFODA	NA		3						
PFBS	NA		44						
PFHxS	NA		56						
PFOS	89.4	12.51	78	46	7	36	7		
PFDS	7.0	12.68	31	36		18	18	9	
PFOSA	NA		19						

Table 3d.   Summary of laboratory performance for PFCs in environmental matrices (unknown solution MXA, results in ng/ml)

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satisfac-tory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% consis-tent LCV	% inconsis-tent LCV
PFBA	59.6	12.52	53	84		16			
PFPeA	29.4	12.54	64	65	22	13			
PFHxA	39.1	12.53	81	86	3	7	3		
PFHpA	24.3	12.55	72	81	12	8			
PFOA	72.4	12.52	97	80	14	3	3		
PFNA	40.8	12.53	78	86	4	7	4		
PFDA	30.0	12.54	83	77	10	3	10		
PFUdA	20.2	12.56	69	84		8	8		
PFDoA	20.0	12.56	67	71	4	21	4		
PFTrDA	18.5	12.57	31	73		18	9		
PFTeDA	18.3	12.57	28	40	20	30	10		
PFHxDA	NA		3						
PFODA	NA		3						
PFBS	36.6	12.53	72	69	19	8	4		
PFHxS	28.5	12.54	75	78	7	11	4		
PFOS	57.8	12.52	97	71	17	9	3		
PFDS	18.5	12.57	33	67		8	25		
PFOSA									

Table 4a.   Summary of laboratory performance for PFCs in human matrices (plasma A, results in ng/ml)

Deter-minand	Assig-ned value	Total error%	% of the data received	% of Z-scores |Z|<2 (Satisfac-tory)	% of Z-scores 3>|Z|>2 (Question-able)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Extreme)	% consis-tent LCV	% inconsis-tent LCV
PFBA	NA		45						
PFPeA	NA		41						
PFHxA	NA		59						
PFHpA	0.2	19.55	77	59	6	6	12	6	12
PFOA	2.9	12.94	95	86	10	5			
PFNA	0.8	14.03	91	55	20	15		10	
PFDA	0.2	18.28	86	53	21	5		11	11
PFUdA	0.1	22.07	59	38	8	8	15	23	
PFDoA	NA		55						
PFTrDA	NA		27						
PFTeDA	NA		27						
PFHxDA	NA		5						
PFODA	NA		5						
PFBS	NA		45						
PFHxS	3.3	12.87	91	75	10	15			
PFOS	10.6	12.62	95	48	29	24			
PFDS	NA		23						

Table 4b.   Summary of laboratory performance for PFCs in human matrices (plasma B, results in ng/ml)

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satis-factory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% con-sistent LCV	% incon-sistent LCV
PFBA	NA		45						
PFPeA	NA		41						
PFHxA	NA		59						
PFHpA	0.2	19.97	77	53	18		12	6	12
PFOA	2.9	12.93	91	90	5	5			
PFNA	0.8	14.00	91	75	15	5		5	
PFDA	0.2	17.76	86	63	11	5		11	11
PFUdA	0.1	21.35	64	57		7	7	21	
PFDoA	NA		50						
PFTrDA	NA		27						
PFTeDA	NA		27						
PFHxDA	NA		5						
PFODA	NA		5						
PFBS	NA		45						
PFHxS	3.3	12.88	91	75	5	20			
PFOS	10.6	12.62	91	60	25	15			
PFDS	NA		23						

Table 4c.   Summary of laboratory performance for PFCs in human matrices (unknown solution MX-3, results in ng/ml)

Deter-minand	Assig-ned value	Total error %	% of the data received	% of Z-scores |Z|<2 (Satis-factory)	% of Z-scores 3>|Z|>2 (Ques-tionable)	% of Z-scores 6>|Z|>3 (Unsatis-factory)	% of Z-scores |Z|>6 (Ex-treme)	% con-sistent LCV	% incon-sistent LCV
PFBA	222.0	12.51	73	69	25	6			
PFPeA	116.2	12.51	68	67	13	20			
PFHxA	144.2	12.51	82	72	11	17			
PFHpA	99.5	12.51	91	80	5	15			
PFOA	251.6	12.50	95	90	5		5		
PFNA	153.4	12.51	91	75	25				
PFDA	118.2	12.51	91	80	5	10	5		
PFUdA	77.7	12.52	86	79	11	11			
PFDoA	71.2	12.52	86	68	16	11	5		
PFTrDA	79.3	12.52	50	82			18		
PFTeDA	84.0	12.51	55	67	8	8	17		
PFHxDA	NA		14						
PFODA	NA		14						
PFBS	137.9	12.51	73	69	6	19			
PFHxS	106.9	12.51	91	90		10			
PFOS	214.5	12.51	95	67	19	10	5		
PFDS	83.9	12.51	55	33	33	8	17		

4 Discussion

Laboratories from Asia, Europe and North-America participated in the present study. For the environmental study, 35 laboratories submitted data on the unknown solution (IVM-PFC-MXA), 36 on the water sample, 27 on the fish sample and 27 on the sludge sample.

4.1 Laboratory performance for different samples

As presented in Table 1 to 4, there was a high variability between matrices. This can also be seen from the relative standard deviations per matrix and per compound in Figure 4 and 5. For the human plasma samples, performance between sample A and B is similar. Because these samples originate from the same plasma batch, results were expected to be similar. Only for PFTrA and PFBS, there is a different performance between both samples, which is explained by the low levels (close to the LOQ) and consequently the limited number of labs that submitted results for these PFCs (n=1 to n=2). The results of PFOA, PFNA, PFDA, PFHxS and PFOS are similar for the plasma samples and the unknown solution. The RSD values for the other PFCs are much larger in the plasma samples as compared to the unknown solution, which is explained by the low levels and number of outlying values (e.g. PFUdA and PFDoA). Satisfactory Z-scores (i.e. Z≤|2|) were obtained by more than 53-90% of the laboratories that submitted results for plasma A and B, except for PFOS and PFUdA in plasma A (see Table 4). The performance for the unknown solution (PFC-MX3) was even better (67-90%, except PFDS).

Concerning the environmental samples, the majority of the laboratories reported data for the unknown solution IVM PFC-MXA that were close to the theoretical values (see Appendix 2). However, surprisingly the RSDs for this sample was worse than for the PFC-MX3 solution, especially for the longer chain acids (RSDs>100%). This is explained by several labs that deviated 50% or more from the theoretical value, and one laboratory (L39) reported values approx. 10-fold higher than the theoretical value. A calculation error is the likely cause of these high values. When these are removed from the dataset, the RSDs drop considerably. A substantial deviation from the theoretical values means that instrument calibration requires attention. The Z-scores were satisfactory for 65-86% of the submitted results (except for PFTeDA, 40%, Table 3).

For the fish sample, the RSDs are worse as compared to the 2008 interlaboratory study (van Leeuwen et al., 2009). In the current study, RSDs vary from 47 to 259%, whereas in the 2008 study, RSDs ranged from 22 to 30% for all compounds except PFOSA (47%). It should be noted that in the 2008 study a very controlled situation was achieved with all labs using multiple mass labels analogues as IS. Also, the data was critically assessed in a meeting and outliers were removed if technical reasons were found. Therefore, the 2008 study represents a situation of 'best possible practice', whereas the current situation represents the current status of intercomparability. Other possible reasons for the current worse performance are (i) several labs only using only mass labelled PFOS and PFOA IS for correction of all compounds; (ii) no outliers were removed during statistical evaluation of the current results (iii) the levels in the earlier study sample were higher (15-23 ng/g for all compounds except PFOS (150 ng/g)), whereas in the current study, levels range from 0.1-2.7 ng/g for all (except PFOS 67 ng/g). This 1-2 order concentration difference also influences the performance of the laboratories. Despite the apparent reasonable RSD of PFOS, it

should be noted that there's a wide distribution of the data with a >40-fold difference between the lowest and highest reported value. This will be discussed in more detail below. Z-scores could be calculated for 6 compounds and generally approx. 50% of the results were satisfactory. 22-46% of the results were unsatisfactory of extreme, showing that there's much room for improvement.

The RSDs for the water sample in most cases RSD values are similar to the earlier study (van Leeuwen *et al.*, 2009), except for PFDA, PFUdA and PFDoA which have much higher RSD values. This is partly caused by the low levels close to the LOQ and by some extremely low values reported by one lab (L04) who reported data 5-6 orders of magnitude lower than other labs, most likely due to a calculation error (see Appendix 2). Also calibration difficulties due to the deliberate pronounced presence of branched PFOS isomers in samples plays a role (as will be explained below). For the other compounds in the water sample, RSDs varied from 35-61%. Some laboratories (2, 7, 34, 35 and 38) reported to have followed the ISO 25101 standard for analysis of PFOS and PFOA in water. Their results were comparable to the other laboratories, showing equal performance in this study. This is somewhat surprising as a more coherent dataset was expected due to the strict ISO25101 standard. In fact, the validation data underlying the ISO25101 standard shows RSD values of 16-36% (Strub, 2008), being only slightly better than observed in the current study. To achieve these lower RSD values in the ISO25101 dataset, substantial amounts of data were discarded as outlying values (Strub, 2008), suggesting limited robustness. Z-scores for the water sample were satisfactory for 27-56% (except PFHxS, see Table 4). The RSDs for the sludge sample are generally below 100%, except for PFHxA, PFHpA, PFNA and PFBS. Although the exact cause of this was not further investigated, this is probably caused by a combination of low levels for these compounds and the lacking use of mass labelled analogue IS for every PFC that is being reported. In addition, many labs may not be very experienced with the analysis of sludge samples. Satisfactory Z-scores were obtained for 24-59% of the laboratories.
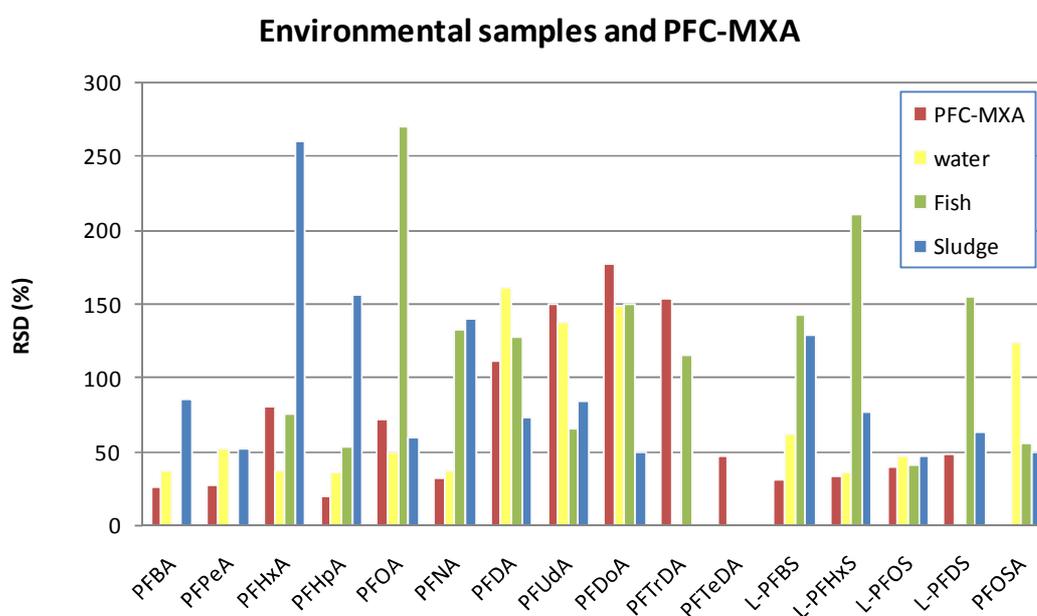

Figure 4   Group performance of all laboratories for different PFCs in the environmental samples and the standard solution

The results for the human plasma and standard solution MX3 which was distributed to all participants in the human study was very good for most of the compounds including PFOS and PFOA. This is illustrated below in Figure 5 where a summary of the results are presented. The RSD, after removing obvious outlier's using Mandel's outlier test, corresponding to removing outliers outside 2-3 x the RDS depending on the number of entries and the distribution of the data. The results for PFOA, PFNA, PFDA, PFHxS and L-PFOS are in the same range as the RSD for the standard solution and below 40%. This shows a surprisingly good agreement between the participating laboratories. For some of the other compounds the RSD was significantly higher, it should however be noted that the levels of these compounds were very low and often below the LOD of most laboratories, resulting both in a limited number of entries and a larger variation between the submitted entries.
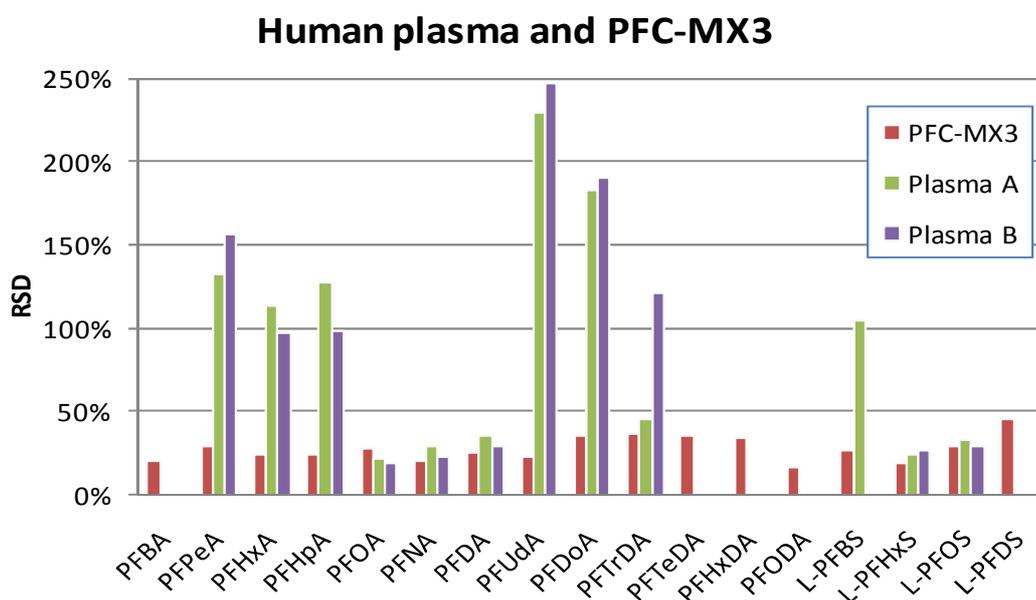
Figure 5   Group performance of all laboratories for different PFCs in the human samples and the standard solution

Some laboratories may have a structural over or underestimation of the results. This can be judged from the grouped Z-score plots as shown in Appendix 4. For example some laboratories had both a negative and positive (resp.) bias for both plasma samples and the unknown solution. Because it includes also the solution, this suggests that the calibration of these laboratories is faulty and should be corrected. The detailed laboratory specific information in the Appendix is thus very useful for the individual participants to improve the quality of the data.

4.2   Possible causes for data variance: the case of PFOS.

There are a number of suggestions that could explain the variance in the PFOS data observed in (especially) the environmental samples. Although the data in Figure 4 suggest that PFOS is among the compounds with lowest RSDs in the environmental samples, it should be noted that the data in the water sample varies from 32 to 180 ng/L (after exclusion of 2 obvious outliers). This is an unsatisfactory wide span of the data. Some experimental variables require more attention than they may have had.

1. Different response factors of branched PFOS isomers

Typically, participants monitor two MS transitions (i.e. 499>99 as qualifier and 499>80 as quantifier) and quantify against a standard consisting of only the linear PFOS isomer (L-PFOS). Because the branched isomers have different response factors compared to the linear isomer (Benskin et al., 2007), a biased result is observed. This problem is limited when the PFOS isomer profile in a sample is dominated by L-PFOS.

However, the water sample showed a pronounced presence of branched isomers because it originated from an AFFF contaminated site. The same holds for the plasma sample. This was demonstrated by laboratories quantifying all PFOS isomers present by m/z 80 and 99 transitions and found relatively high concentrations (180 and 123 ng/L) using the different transitions. In this case, it would have been more appropriate to quantify against a standard containing also the branched isomers in a similar composition. However such calibration solution might still be hard to find. Isomer specific quantitative standards have recently become available but were not available at the time of this study to calculated isomer specific relative response factors. The optimal calibration method for PFOS to be used may vary case by case. Therefore, chemists should critically assess the isomer profile of the sample in order to determine the best calibration method.

2. Reporting results on the anion basis

Another source of variance is the calibration based on the anion or the salt. PFSA anions have a cationic counter ion. Common counter ions are K^+, Na^+ and NH_4^+. Commercial standards for PFOS are sold with either one of these counter ions (e.g. K^+ for the Wellington standard and NH_4^+ for the Fluka standard). These different counter ions result in different molecular masses (MW (g/mol) 538 for PFOS-K, 522 for PFOS-Na and 517 for PFOS-NH_4) and this will lead to different concentration of PFOS anions in a calibration solution, if not accounted for. Correction factors are provided in the Table 5 below.

This problem is even more pronounced for PFHxS and PFBS because the counter ion mass becomes larger relative to the PFSA anion. It is therefore strongly recommended to report results for the PFSAs based on the anion (correction factors see Table 5). In this study, this was also recommended, but nevertheless several laboratories have not followed up on this, either because it was not part of their routine, or due to a misunderstanding in communication and instruction.

It is however important that harmonisation is more strictly applied. ISO 25101 specifies that results should be reported on an anion basis. In the future only concentrations based on the anion are accepted in QA/QC studies and when reporting other results of the PFSAs which dissociate.

Table 5   Correction factors calculated for some PFSA salts

Compound	MW anion (g/mol)	MW including cation (i.e. salt) (g/mol)	Correction factor
PFBS-K	299	338	0.88
PFHxS-Na	399	422	0.95
PFHpS-Na	449	472	0.95
PFOS-K	499	538	0.93
PFOS-Na	499	522	0.96
PFOS-NH4	499	517	0.97
PFDS-Na	599	622	0.96

Both the isomer distribution and the different reporting of the PFSAs can influence the results and although the effect might be of minor influence, the assigned value and the resulting Z-scores for PFOS in the environmental samples should be interpreted with some care.

5   Conclusions

A large number of laboratories participated in this international interlaboratory study, showing that the interest in analysis of PFCs is growing as well as the intention to evaluate laboratory performance so as to improve the data produced in the field. This study showed that the performance of labs participating in the human part of the study was better than the performance in the environmental part. One of the reasons may be that the overall experience in that study is larger as compared to the environmental part. Sources that have contributed the variance in this study are:

- the limited use of mass labelled internal standards. Several laboratories only used one or two, whereas it is strongly recommended to use a mass labelled analogue for every single PFC that is being analysed and quantified.

- labs reporting results of the PFSAs on the basis of the salt rather than on the basis of the anion. The latter is strongly recommended.

- the pronounced profile of branched PFOS isomers in the water sample which complicated an accurate calibration. It is recommended to judge case by case whether calibration should be performed with a standard with only linear isomer, or with a standard containing linear and branched isomers.

- the concentrations of several PFCs in the environmental samples were lower (close to the LOQ) than in earlier interlaboratory studies. At these levels, performance of methods becomes less accurate and precise, which is reflected in a higher variance.

- the methods used for blank correction are unclear and especially when levels were very low, this could have affected the results and RSD of the compounds close or below the LOD reported by the participants.

- the methods for correction for recovery were not reported and seem to differ between the laboratories, some laboratories did not correct for recovery at all. The effect of this on the total results is unclear.

All in all, the results show that there is room for improvement, especially for the labs that participated in the environmental part of this study.

Acknowledgements

References

Benskin, J.P., Bataineh, M. & Martin, J.W. (2007). Simultaneous characterization of perfluoroalkyl carboxylate, sulfonate, and sulfonamide isomers by liquid chromatography-tandem mass spectrometry. *Analytical chemistry, 79*(17), 6455.

de Boer, J. (1997). The preparation of biological reference materials for use in inter-laboratory studies on the analysis of chlorobiphenyls, organochlorine pesticides and trace metals *Marine Pollution Bulletin, 35*(1-6), 84-92.

Cofino, W.P., Wells, D.E., Ariese, F., van Stokkum, I., Wengener, J.W. & Peerboom, R. (2000). *J. Chemometrics and Intelligent Laboratory Systems, 53,* 37-55.

Cofino, W. P., van Stokkum, I.H.M., van Steenwijk, J. & Wells, DE. (2005). A new model for the inference of population characteristics from experimental data using uncertainties. Part II. Application to censored datasets. *Analytica Chimica Acta, 533*, 31–39.

de Vijver, K.I.V., Hoff, P.T., van Dongen, W., Esmans, E.L., Blust, R. & de Coen, W.M. (2003). Exposure patterns of perfluorooctane sulfonate in aquatic invertebrates from the Western Scheldt estuary and the southern North Sea. *Environ. Toxicol. Chem., 22*(9), 2037-2041.

ISO 13528:2005. Statistical methods for use in proficiency testing by interlaboratory comparisons.

ISO 25101:2009 Water quality — Determination of perfluorooctanesulfonate (PFOS) and perfluorooctanate (PFOA) — Method for unfiltered samples using solid phase extraction and liquid chromatography/mass spectrometry.

ISO/IEC 17043:2010 Conformity assessment — General requirements for proficiency testing.

Lindström, G., Kärrman, G. & van Bavel, B. (2009). Accuracy and precision in the determination of perfluorinated chemicals in human blood verified by interlaboratory comparisons *Journal of Chromatography A, 1216*(3), 394-400.

Martin, J.W., Kannan, K., Berger, U., de Voogt, P., Field, J., Franklin, J., Giesy, J. P., Harner, T., Muir, D.C.G., Scott, B., Kaiser, M., Jarnberg, U., Jones, K.C., Mabury, S.A., Schroeder, H., Simcik, M., Sottani, C., van Bavel, B., Karrman, A., Lindstrom, G. & van Leeuwen, S. (2004). Analytical challenges hamper perfluoroalky research. *Environmental Science & Technology, 38*(13), 248A-255A.

Smithwick, M., Norstrom, R.J., Mabury, S.A., Solomon, K., Evans, T.J., Stirling, I., Taylor, M.K. & Muir, D.C.G. (2006). Temporal trends of perfluoroalkyl contaminants in polar bears (Ursus maritimus) from two locations in the North American Arctic, 1972-2002. *Environmental Science & Technology, 40*(4), 1139-1143.

So, M.K., Taniyasu, S., Yamashita, N., Giesy, J. P., Zheng, J., Fang, Z., Im, S.H. & Lam, P.K.S. (2004). Perfluorinated compounds in coastal waters of Hong Kong, South China, and Korea. *Environmental Science & Technology, 38*(15), 4056-4063.

Strub, M.P. (2008). *Interlaboratory comparison 08/95314 – Validation of ISO/DIS 25101.* Report, INERIS, France.

Thompson, M. & Wood, R. (1993). International Harmonised Protocol for Proficiency Testing of (Chemical) Analytical Laboratories. *Journal of AOAC International, 76*, 926-940.

Thompson, M., Ellison, S.L.R. & Wood, R. (2006). The International Harmonised Protocol for Proficiency Testing of Analytical Chemistry Laboratories. IUPAC Technical Report. Interdivisional Working Party for Harmonization of Quality Assurance Schemes. *Pure Appl. Chem., 78,* 145–196.

van Leeuwen, S.P.J., Karrman, A., van Bavel, B., de Boer, J. & Lindstrom, G. (2006). Struggle for quality in determination of perfluorinated contaminants in environmental and human samples. *Environmental Science & Technology*, *40*(24), 7854-7860.

van Leeuwen, S.P.J., Swart, C.P., van der Veen, I. & de Boer, J. (2009). Significant improvements in the analysis of perfluorinated compounds in water and fish: Results from an interlaboratory method evaluation study. *Journal of Chromatography A, 1216*(3), 401-409.

Wells, D.E., Cofino, W.P. & Scurfield, J. A. (2004). *The Application of the Cofino Model to Evaluate Laboratory Performance Study Data using the BandWidth Estimator.* FRS Marine Laboratory, Aberdeen, Collaborative Report No. 04/04.

Wells, D.E. & Scurfield, J.A. (2004). *Assessment Rules for the evaluation of the QUASIMEME Laboratory Performance Studies Data* – version 2, February 2004. QUASIMEME Project, FRS Marine Laboratory, 375 Victoria Road, Aberdeen AB11 9DB.

Appendices

1. List of participants.
2. Results and graphical representation
3. Numerical Z-score values per matrix
4. Z-score plots per compound per matrix.
5. Additional method information.