UN environment programme

Organization and Outcomes of Four Rounds of Interlaboratory Assessments on Persistent Organic Pollutants

ISBN: 978-92-807-4101-8
Job number: DTI/2591/GE

Suggested citation: United Nations Environment Programme (2023). *Organization and Outcomes of Four Interlaboratory Assessments on Persistent Organic Pollutants.* Geneva.

Production: United Nations Environment Programme

URL: https://wedocs.unep.org/20.500.11822/44155
DOI: https://doi.org/10.59117/20.500.11822/44155

Acknowledgements

This report was developed in the framework of the projects titled “Implementation of the POPs Monitoring Plan in the Asian Region" and "Continuing regional Support for the POPs Global Monitoring Plan under the Stockholm Convention in the"- Africa, Pacific and Latin-American and Caribbean regions, funded by the Global Environment Facility; Project GEF-ID 4894, GEF ID 6978, GEF ID 4881, and GEF ID 4886, and was coordinated by Vrije Univeriteit (VU), Netherlands and Man-Technology-Environment (MTM) research center, Örebro University, Sweden.

This document has been prepared by:

Dr. Heidelore Fiedler, Örebro University, School of Science and Technology for Basel Convention Coordinating Center Stockholm Regional Center Latin America and the Caribbean.

Internal review at UNEP was conducted by Victor Hugo Estellano, Haosong Jiao, Tapiwa Nxele and Zhanyun Wang.

Layout and graphic design: Mireia Mas Vivancos and Lowil Espada.
All graphs and tables were prepared by Heidelore Fiedler unless otherwise specified.

UNEP, Chemicals and Health Branch, Geneva, Switzerland. 74 pp

Abbreviations

AV	Assigned value
CEE	Central and Eastern Europe
DDT	Dichlorodiphenyltrichloroethane
dl-PCB	Dioxin-like polychlorinated biphenyls
dl-POPs	Dioxin-like persistent organic pollutants Include: 29 congeners that were assigned a TEF by WHO/IPCS expert group, namely polychlorinated dibenzo-para-dioxins (7), polychlorinated dibenzofurans (10), and polychlorinated biphenyls (12)
GC	Gas chromatograph(y)
GC/ECD	Gas chromatography - electron capture detection
GC/MS	Gas chromatography - mass spectrometry
GEF	Global Environment Facility
GRULAC	Group of Latin America and the Caribbean
HBCD	Hexabromocyclododecane
HCB	Hexachlorobenzene
HCH	Hexachlorocyclohexane
HRGC	High resolution gas chromatograph or high resolution gas chromatography
HRMS	High resolution mass spectrometer or high resolution mass spectrometry
ISO	International Standardization Organization
ILAC	International Laboratory Accreditation Cooperation
LB	Lower-bound
LRMS	Low resolution mass spectrometry or low resolution mass spectrometer
MS	Mass spectrometer or mass spectrometry
MTM	Man-Technology-Environment
ND	Not detected
OCP	Organochlorine pesticide
OECD	Organisation for Economic Co-operation and Development
PBDE	Polybrominated diphenyl ethers
PCB	Polychlorinated biphenyls
PCDD/PCDF	Polychlorinated dibenzo-*para*-dioxins/polychlorinated dibenzofurans
PFAS	Per- and polyfluoroalkyl substances
POPs	Persistent organic pollutants

QUASIMEME	Quality Assurance of Information for Marine Environmental Monitoring in Europe
QA/QC	Quality assurance/quality control
TEF	Toxicity equivalency factor
TEQ	Toxicity equivalent
TEQ_{PCB}	Toxicity equivalent based on dl-PCB
$TEQ_{PCDD/PCDF}$	Toxicity equivalent based on PCDD and PCDF (dl-PCB not included)
TEQ_{total}	Toxicity equivalent based on PCDD, PCDF, and dl-PCB
UB	Upper-bound
UNEP	United Nations Environment Programme
VU E&H	Environment & Health of the Vrije Universiteit Amsterdam
WEOG	Western European and Other Groups

Table of contents

List of figures

List of tables

List of figures in appendix

List of tables in appendix

1

Introduction

Persistent organic pollutants (POPs) have been shown to adversely impact human health and the environment. Concern over POPs is also attributed to their stability and persistence in the environment, potential to undergo long-range transport, and to accumulate in animals, humans and food chains. Men and women differ in their physiological susceptibility to the effects of exposure to POPs and exposure can be impacted by societal and occupation roles (United Nations Environment Programme [UNEP] 2019). The objective of the Stockholm Convention (SC) on POPs is to protect human health and the environment from POPs with the ultimate goal to eliminate them, where feasible. Accurately measuring and analyzing of the concentrations POPs is an important step towards evaluating the effectiveness of the convention and the potential impacts of POPs in human and the environment (UNEP 2021a).

The interlaboratory assessment accompanies the United Nations Environment Programme's (UNEP) capacity building programme for laboratories analysing POPs. The programme implements the recommendations of the Conference of the Parties (COP) to the Stockholm Convention as expressed in the Guidance on the global monitoring plan for POPs (hereinafter referred to as the guidance document) in article 16 of the Convention (UNEP 2021a). In chapter 4, the guidance document states that "interlaboratory exercises are often used to assess the effectiveness of quality assurance/quality control (QA/QC) practices among several participating labs and to provide a measure of interlaboratory comparability. This usually involves the circulation and analysis of a common standard or reference sample, often at two or more concentration levels" (UNEP 2005). In order to determine the 'true' concentration of (in this case) POPs in a sample, a chemical laboratory must be able to prove that it is able to identify and quantify chemicals (analytes) of interest at concentrations of interest. Such accuracy and precision in the determination of POPs is required by article 16 of the Convention and subsequent guidance developed for the global monitoring plan (GMP). The needs and support are documented in COP decisions SC-3/16, SC-4/31, SC-5/18 and SC-6/23, and in chapter 3 of the guidance document. To provide reliable monitoring information for the Parties to the Stockholm Convention, the guidance document aims to detect a 50% decrease in the levels of POPs within a 10 year period (UNEP 2021a). This means that POPs laboratories must be capable – at any time – to analyse samples for POPs within a margin of ±25% (Abalos *et al.* 2013).

In an interlaboratory assessment, all participating laboratories analyse the same sample within a limited time frame for previously determined analytes and report the results to the coordinator of the study. All results are evaluated together according to international standards, as established by the International Standardization Organization (ISO) or the International Laboratory Accreditation Cooperation (ILAC), thus allowing a performance classification.

Where proficiency tests or 'round robins' on polychlorinated biphenyls (PCB), organochlorine pesticides (OCPs), and dioxin-like (dl-)POPs are well established for laboratories in many OECD countries, challenges can be expected for developing country laboratories. This is also the case in some OECD countries as well, since they do not yet all have the necessary experience to analyse a large number of POPs in biotic and abiotic matrices at the requested accuracy and within the required time limits.

To assist laboratories in improving the quality of their analyses, UNEP has organized regional capacity building and training programmes, which started in 2009. As part of this activity, the first round of the Global Interlaboratory Assessment on POPs was organized in

2010-2011 (UNEP 2012; Abalos *et al.* 2013; van Leeuwen *et al.* 2013), the second in 2012-2013 (van Bavel *et al.* 2014; UNEP 2015; Fiedler, de Boer and van Bavel 2016). This third round was implemented in 2016-2017 (UNEP 2017; van der Veen, de Boer and Fiedler 2017) and the fourth in 2018-2019 (UNEP 2021b). Further information is contained in published papers in the context of the UNEP-coordinated interlaboratory assessments such as in a general context de Boer *et al.* (2008), POP-specific assessments as for PFAS (Fiedler, van der Veen and de Boer 2020; van der Veen, Fiedler, and de Boer 2023), dioxin-like compounds (Fiedler, van der Veen and de Boer 2022a), organochlorine pesticides and brominated flame retardants (de Boer, van der Veen and Fiedler 2022), or on countries supported by UNEP (Fiedler, van der Veen and de Boer 2022b).

The "Report on International Intercalibration Studies" (UNEP 2005) emphasizes the importance of accurate results in POPs analysis, with an analytical variance to be as small as possible in order to make data acceptable and comparable between laboratories, countries, and regions, so as to allow sound decision making. Participation at international intercalibration assessments is considered a prerequisite for existing and well established as well as for newly set-up laboratories because there is a need to permanently check the laboratory's performance and 'prove' their capabilities. From an international quality assurance point of view, world-wide international studies are preferred, but national initiatives could also improve the analytical quality in a country or a region.

Within the framework of UNEP's capacity building project for training of laboratory staff on POPs analysis in developing countries, the Department of Environment & Health of the Vrije Universiteit Amsterdam, the Netherlands (VU E&H) and the Man-Technology-Environment (MTM) Research Center, School of Science and Technology at the University of Örebro, Sweden, have organised four rounds of the Bi-ennial Global Interlaboratory Assessment on Persistent Organic Pollutants (POPs). A summary of the organization, the assessment approach and the performance of the laboratories are compiled in this report.

©UNEP/Elsemieke de Boer

2

Materials and methods

The number and type of test samples have increased as new POPs were listed in the Annexes of the Stockholm Convention. At present, the POPs studied include polychlorinated dibenzo-p-dioxins (PCDD), polychlorinated dibenzofurans (PCDF), polychlorinated biphenyls (PCB) and organochlorine pesticides (OCPs), *i.e.*, DDT and metabolites, aldrin, dieldrin, endrin, chlordane, hexachlorobenzene (HCB), heptachlor and *cis*-heptachlorepoxide, and mirex. The 'new´ POPs, polybrominated diphenylethers (PBDE), hexachlorocyclohexanes (HCHs), chlordecone (kepone), pentachlorobenzene (PeCBz), α- and β-endosulfan, endosulfan sulphate and perfluorinated alkylsulphonates (PFSA) as well as hexachlorobutadiene (HCBD). A separate test solution was prepared, and an own assessment was done for toxaphene (three Parlar congeners) in the 4th round. Hexabromobiphenyl (HxBB) (as polybrominated biphenyl, PBB 153) was provided as a single compound in Round 3 (IL3); in IL4, PBB 153 was included in the PBDE test solution.

In the last round, 16 matrices were offered for analysis: nine test solutions to cover all POPs, two air extracts (one in toluene for the chlorinated and brominated POPs and one in methanol for the fluorinated POPs), sediment, fish, human milk, human plasma and water (the latter two for PFAS only). The test solutions were ampouled in amber glass ampoules with the target compounds in undisclosed concentrations. The air extracts were also ampouled, sediment was air-dried, the fish (crab) was sterilized in glass jars, the plasma frozen and the human milk was homogenized, frozen and stored at -20°C prior to shipment. Water was sent in high-density polyethylene (HDPE) bottles.

2.1 Identification and preparation of the test samples

There were two broad groups of test samples: (a) test solution of native POPs in an inert solution and (b) naturally contaminated environmental or human matrices (some of them were amended by a given analyte since the contamination was too low for quantification).

2.1.1 Test solutions of analytical standards

All analytes of the POPs listed into either Annexes A, B, or C of the Stockholm Convention were provided in mixtures of test solutions consisting of POPs that can be analyzed by similar instrumentation in an organic solvent. Table 1 provides an overview on the test solutions sent to laboratories in the four rounds of the international assessments.

2.1.2 Naturally contaminated test samples

Table 2 provides an overview of the naturally contaminated test samples. It should be mentioned that the air extracts (IL2-IL4), and the 'fish toxaphene' (IL4) were amended with native POPs since the environmental samples collected had most of the POPs below the limit of quantification.

Table 1: Test solutions of analytical standards for the analysis of POPs in the four rounds

	1st Round	2nd Round	3rd Round	4th Round
Test solutions: Prepared, ampouled and labelled by • **E&H VU Amsterdam from crystals obtained from Da Vinci Laboratory Solutions B.V. (Rotterdam, The Netherlands)** • **Cambridge Isotope Laboratories, Inc. (Tewksbury, USA)** • **Wellington Laboratories (Guelph, Canada)**				
OCPs	The **test solution** in *iso*-octane with the following OCPs present: aldrin, dieldrin, endrin, *cis*-chlordane (alpha), *trans*-chlordane (gamma), oxychlordane, *cis*-nonachlor, *trans*-nonachlor, heptachlor, *cis*-heptachloroepoxide, *trans*-heptachloroepoxide, *o,p'*-DDT, *p,p'*-DDT, *o,p'*-DDD, *p,p'*-DDD, *o,p'*-DDE, *p,p'*-DDE, hexachlorobenzene, mirex		In addition: α-HCH, β-HCH, γ-HCH, α-endosulfan, β-endosulfan, endosulfan sulfate, chlordecone, hexachlorobutadiene, and pentachlorobenzene	
PCB(6)	Mixture of the indicator PCB (six congeners) in *iso*-octane			
PBDE and PBB	Not contained	A mixture of eight PBDE and separately PBB 153 in nonane	A mixture of eight PBDE and separately PBB 153 in nonane	A mixture of nine PBDE and PBB153 in *iso*-octane
Toxaphene	Not contained			A mixture of Parlar 26, 50, 62 in nonane prepared by VU E&H out of individual stock solutions
PCDD/PCDF	A mixture of seventeen 2,3,7,8-substituted PCDD/PCDF congeners in nonane			
dl-PCB	A mixture of twelve dl-PCB in nonane			
HBCD	Not contained	Not contained	A mixture of α-, β-, and γ-isomers in toluene	
PFAS	Not contained	A mixture of PFAS (PFCA[1], PFSA[2] and FOSA[3]) in methanol. A mixture of PFOS precursors (MeFOSE4,(EtFOSE)5, MeFOSA6, and EtFOSA) 7	A mixture of PFAS with perfluoroalkyl carboxylic acids (PFCA) and perfluoroalkane sulfonic acids (PFSA), perfluorooctane sulfonamides (FOSAs) and perfluorooctane sulfonamidoethanols (FOSEs)	

Table 2: Test samples for the analysis of POPs in the four rounds – Naturally contaminated test samples

	1st Round	2nd Round	3rd Round	4th Round
Abiotic test samples				
Sediment	Samples were dried at 40°C, sieved (at 0.5 mm), homogenized, filled into plastic containers, and stored at room temperature until shipment. Samples were from WEPAL8			
		A marine sediment from the Netherlands	A sediment from the Elbe River, Germany	A sediment from the harbour of Rotterdam, The Netherlands
Air extract	The air test samples were extracts from PUFs in active air samplers taken at different locations. For organochlorine and organobromine POPs analyses, the PUFs were conditioned and extracted with toluene (TOL); for PFAS analysis, the PUFs were conditioned and extracted with methanol (MeOH). Extracts were spiked with native target analytes. The extracts were ampouled into 1.2 mL amber glass ampoules before shipment			
	As an approximation, a fly ash sample from a waste incinerator in Sweden was taken to be analysed for PCDD/PCDF	A toluene extract of polyurethane foam (PUF) taken near one of Sweden's largest hazardous waste incinerators	Air (TOL) for Br-/Cl-analytes) were extracts from PUFs exposed in Barcelona, Spain; amended with native OCPs, PBDE and HBCD). Air (MeOH) was of the same location and amended with PFAS and PFOS precursors	Air (TOL) for Br-/Cl-analytes) were extracts from PUFs exposed in Brno, Czech Republic and in Örebro, Sweden, to which remaining spiked samples (OCPs, PBDE and HBCD) extracts from the 3rd round of the interlaboratory assessment were added); Air (MeOH) was of the same locations and amended with remaining spiked extracts from the 3rd round of the interlaboratory
Water	The water test materials consisted of surface water samples. After bottling in high-density polyethylene bottles (250 mL), the material was sterilized by irradiation.			
	Not contained	Amsterdam harbour, the Netherlands	Pooled, from different locations in the Netherlands	Pooled, from different locations in the Netherlands

	1st Round	2nd Round	3rd Round	4th Round
Biotic test samples				
Fish	After filletting and homogenizing, individual glass screw cap jars were filled with ca. 50 g homogenate. The jars were sterilized by autoclaving, thus, could be stored and transported at room temperature before opening			
		A pike-perch filet from the Netherlands	Chinese mitten crab from the Netherlands	'Fish A' a pike perch from river Amer from the Netherlands. 'Fish toxaphene' was the same pike perch fortified with toxaphenes
Human milk	The test material consisted of pooled homogenized human milk from milk banks in Sweden. 50 mL milk was packed in polypropylene bottles and frozen prior to shipment			
	Human milk from Swedish mothers	Human milk from the Örebro region	Human milk from the Örebro region	A pooled human milk sample from four milk banks in Sweden; amended with cows' milk from Sweden (approx. 25%; to reach the sample volume necessary for this interlaboratory assessment)
Human blood	The human blood samples consisted of pooled human plasma from the general population and people occupationally exposed to PFAS. 1 mL of homogenized sample was placed in a polypropylene vial and kept frozen until shipment			
	Not contained	Occupationally exposed people were professional ski wax technicians	Occupationally exposed people were firefighters	Occupationally exposed people were firefighters

[1] PFCA	Perfluoroalkane carboxylic acids	5 EtFOSE	N-ethyl perfluorooctane sulfonamidoethano
[2] PFSA	Perfluoralkane sulfonic acids	6 MeFOSA	N-methyl perfluorooctane sulfonamide
[3] FOSA	Perfluorooctane sulfonamide	7 EtFOSA:	N-ethyl perfluorooctane sulfonamide
4 MeFOSE	N-methyl perfluorooctane sulfonamidoethanol	8 WEPAL	Wageningen Evaluating Programmes for Analytical Laboratories

2.2 Processing of samples and results

2.2.1 Distribution of test samples

The human milk, human plasma, and the air extracts as well as the test solutions for PCDD/PCDF, dl-PCB, HBCD, and PFAS were distributed by MTM Research Centre. The sediment, fish, and water and the test solutions for OCPs, PCB, PBDE, and toxaphene were distributed by VU E&H. All shipments containing human milk or plasma samples were packed in a polystyrene container with frozen plastic ice blocks.

Each shipment was accompanied by (a) a letter listing the type of test samples contained in the shipment, (b) a customs letter stating the context of the interlaboratory assessment, especially the technical nature and non-commercial approach, and (c) certificates on non-infectiousness of the materials for the human milk and the human plasma. Instructions on the nature of the test materials as well as a file (MsExcel®) to report the results were sent by e-mail to all laboratories.

2.2.2 Reporting results

For each round, all results were combined into one results database (MsExcel®) according to laboratory (laboratory code), analyte and test sample. In this assessment, these aggregated data were shared with the participating laboratories for a confirmation of their data and in addition, laboratories were allowed to make small corrections for obvious errors, such as units, sum parameters, treatment of non-detects, use of decimals.

2.3 Methods used by participants

All participating laboratories used in-house methods for sample preparation, clean-up, extraction and instrumental analysis. It shall be noted that not all laboratories provided information on their methods according to the reporting format. In addition, the definition of "high resolution mass spectrometer" was not interpreted by all laboratories in the same way; here, "HRMS" are understood to be sector-field instruments.

The methods used included modified or adapted standard methods including for example EPA 1613 and EU 1948 for the dl-POP analysis. For PCDD/PCDF and dl-PCB, most laboratories reported that high resolution GC/MS (HRGC/HRMS) systems were used – with the limitations mentioned above. Three laboratories, used to analyse PCB, applied GC/ECD instrumentation for the analysis of dioxin-like PCB and reported on toxic equivalents; they did not analyse PCDD/PCDF.

For the separation of dl-POPs, the most common length for GC columns is still 60 m; only for a few instances, shorter columns – 30 m – were used. Only one laboratory reported using a 50 m column. All participants used an LC-MS/MS method for the analyses of PFAS, and only one reported using a GC method for the analyses of the PFOS precursors.

In the other compound classes, this is more diversified and GC/ECD, low resolution GC/MS (including GCxGC/MS), and HRGC/HRMS were also used.

Sample extraction was performed using variety of techniques and methods. For the extraction, Soxhlet extraction was still the most popular extraction method, although more and more laboratories used accelerated or pressurized liquid extraction that has become more popular.

Several organic solvents such as toluene, hexane, acetone or dichloromethane were used in different combinations for extraction of especially the fish and sediment sample. Of those, a mixture of hexane and acetone was the most preferred combination for the analyses of OCPs and PCBs. For PBDE this combination was also used for fish and sediment, but the most preferred solvent for the sediment sample was toluene. For the extraction of PFAS almost all participants used methanol.

Furthermore, a wide variety of sample clean up open column chromatography was used where acid or base loaded silica was most commonly used followed by Florisil and alumina (especially for the OCPs). For the analysis of dioxins, the majority of the laboratories included a carbon column as the final separation step in agreement with the standard methods. Gel permeation chromatography (GPC) was used by only a few laboratories. Activated copper was often used as an extra clean-up for the sediment sample.

The participants were encouraged to use appropriate GC columns for the analysis, preferably dual-column sets. Although several co-elution issues are known, especially when using ECD as the final detection technique, only few laboratories reported that two columns or a confirmation column was used. This was also true for PCDD/PCDF analysis, where the

use of a confirmation column is described in most official methods; however, this was hardly used by the participating laboratories. The major reason may be that only 2,3,7,8-substituted congeners or dl-PCB were to be reported. In addition, the human milk sample is known to have only the 2,3,7,8-substituted PCDD and PCDF present and thus, there is no need to separate these congeners from more unpolar non-TEF congeners. The other important reason is that custom-made HRGC columns are available for dl-POPs. Only one laboratory used a more sophisticated GCxGC arrangement.

The methodology for the PBDE analysis is similar to that of the OCPs and PCB. The clean-up and extraction are similar and also the final analysis is performed on similar instrumentation, including HR- or LRGC/MS systems.

The sample extraction, clean-up and detection of the more polar PFAS, *e.g.*, PFCA and PFSA, including PFOS, is completely different from the traditional POPs. From the 29 laboratories that submitted results for PFAS, only one laboratory used a time-of-flight instrument; all others reported to use LC-MS/MS. For the separation of the analytes, the majority used HPLC columns; however, also UPLC columns are in use. Normally, a C18 based column was used; but some also used C8-based columns. One laboratory reported to have applied GC/LRMS (using a DB-WAX column, 30 m x 0.25 mm x 0.25 um) for the separation of PFOS precursors, *e.g.*, Me/EtFOSA and Me/EtFOSE.

2.4 Assessment of performance

All participating laboratories were provided with instructions and a template to report results for each of the POP groups electronically (MsExcel®). The laboratories were asked to use their own methods. The approach may result in somewhat more variation but avoids systematic errors that could be introduced when describing a standard method for all participants. 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 of the assessment is based on the standard ISO 13528 (2005) and the International Union of Pure and Applied Chemistry International Harmonised Protocol for Proficiency Testing by Thompson, Ellison and Wood (2006). As for the first round of the Global Interlaboratory Assessment on POPs (UNEP 2012), the performance was assessed according to the QUASIMEME proficiency testing organisation. The assigned value, the between-lab CV values and the laboratory assessment using z-scores are based on the Cofino model (Molenaar, Cofino and Torfs 2018) according to the principles employed in the Quality Assurance of Information for Marine Environmental Monitoring in Europe (QUASIMEME) proficiency testing. The following equation and definitions apply:

THE FORMULA USED IS

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

The z-scores can be interpreted as follows:

$\|z\| < 2$	Satisfactory performance	S
$2 < \|z\| < 3$	Questionable performance	Q
$\|z\| > 3$	Unsatisfactory performance	U
$\|z\| > 6$	Extreme performance	U

Here an assigned value (AV) is considered reliable and statistically valid when the below criteria are met (see section 2.5).

2.5 Criteria for the statistical assessments

Four different categories are used:

Category 1: For data where the number of numerical observations is ≥ 7:

- An AV is based on the mean when ≥ 25% of values have a z-score of $|z| < 2$.
- Where < 25% of the data have $|z| < 2$, the value is indicative, i.e., at least 25% must be in good agreement.

Category 2: For data where the number of numerical observations is > 3 and < 7:

- An AV 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 data sets, 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 where the number of numerical observations is < 4:

- No AV is given. Normally, the median value is given as an indicative value.

Category 4: For data where the high total error > 100% in combination with bad performance, no AV is given.

Since it is not possible to calculate a z-score for values below the limit of detection (LOD), the so-called 'left censored values' (LCVs) are used (Cofino *et al.* 2005) with the following quality criteria:

- LCV/2 < (concentration corresponding to |z|=3): LCV consistent (labelled as 'C') with AV.
- LCV/2 > (concentration corresponding to |z|=3): LCV inconsistent (labelled as 'I') with AV, *i.e.*, LCV reported by laboratory much higher than numerical values reported by other laboratories.

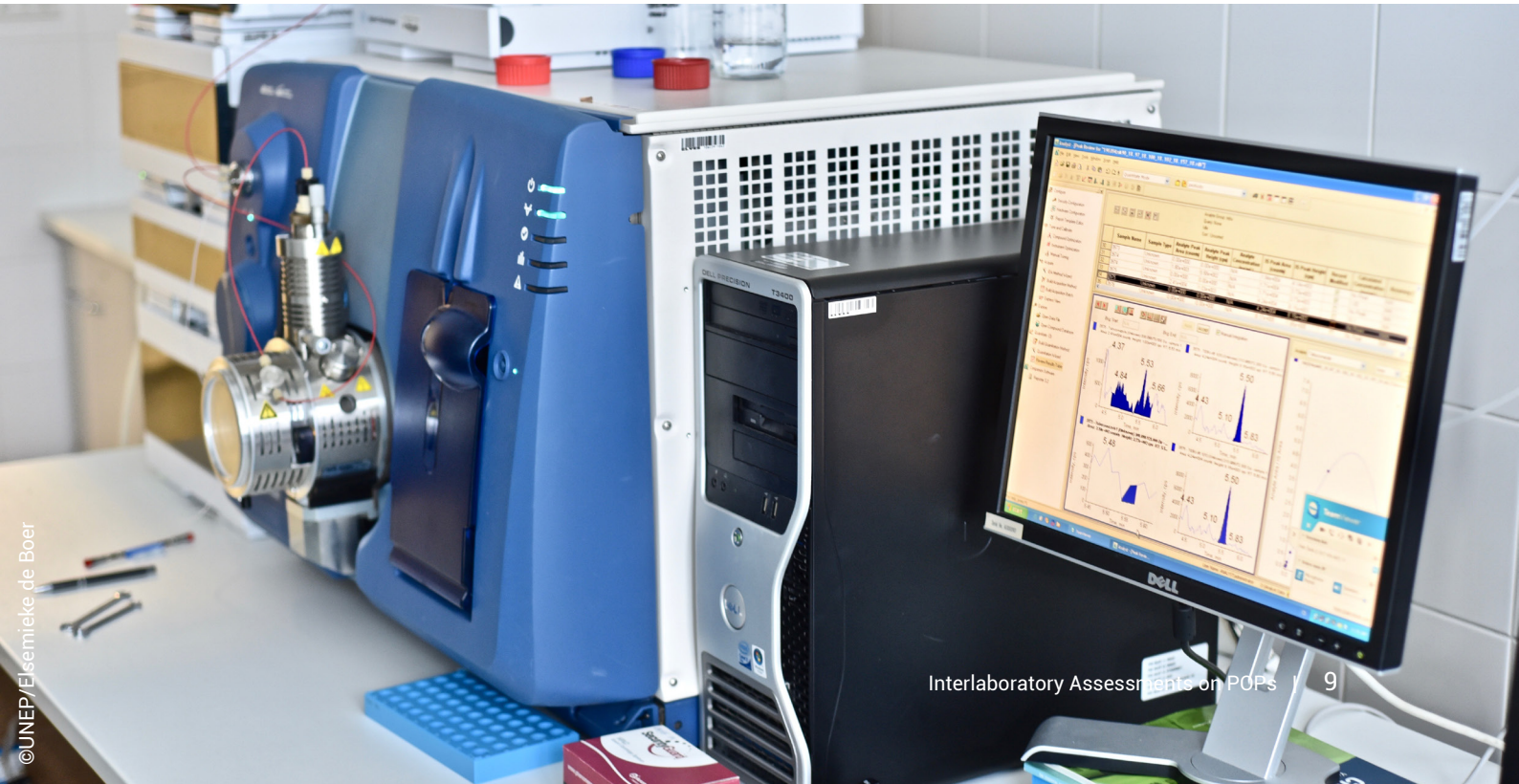
Table 3: Keys and color codes used for the interpretation of z-scores given:

Z-score key:	S – Satisfactory	S
	Q – Questionable	Q
	U – Unsatisfactory	U
LCV key:	C – Consistent	C
	I – Inconsistent	I

It is important to note that, in contrast with many other interlaboratory exercises, but in line with the requirement from the global monitoring plan (GMP) of the Stockholm Convention, all laboratories producing results for the GMP of the Stockholm Convention should be able to distinguish between two values differing 50% from each other. Consequently, a target error of 25% has been set on which the z-scores are based.

2.6 Presentation of results

An example of the visualization of the statistics is shown in Figure 1. With PBDE 99 in sediment as an example, the results were visualized by showing the normal distribution of a given POP determinand in its matrix (upper left), the overlap matrix as kilt plot (upper right), a ranked overview of results and the standard deviations (lower left), and the z-score plot (of all results in the specific round) (lower right).



©UNEP/Elsemieke de Boer

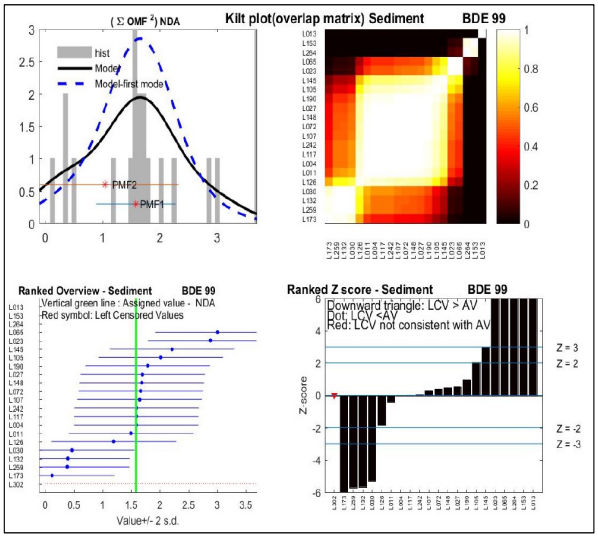


Figure 1: Visual statistics for PBDE 99 in sediment (IL4) (source: de Boer, van der Veen and Fiedler 2022)

2.7 Samples prepared and distributed

The test samples that were prepared by the coordinating laboratories are summarized in Table 1 for the test solutions of analytical standards and test matrices in Table 2. For the first round, the number of test samples prepared and shipped to participating laboratories was not recorded but may be at least 500 samples. For the other three rounds, the number of samples in each round are available as shown in Table 4. The most frequently requested test solutions were for OCPs and indicator PCB with roughly 300 samples, followed by dl-PCB (194 samples), PCDD/PCDF (176 samples) and PBDE (169 samples). Among the test samples, the most frequently requested test sample were fish (414 samples), human milk (328 samples), sediment (321 samples), and air (296). More than three thousand samples have been prepared for the three rounds of interlaboratory assessment. With a 'typical value' of USD 200 for a test solution of analytical standards and USD 700 for test samples, the economic value for the three rounds (IL2-IL4) is approximately USD 1.4 million. The number of samples prepared and distributed in each round is visualized graphically in Figure 2.

Table 4: Number of test samples prepared for participating laboratories (IL2-IL4)

Test solutions of analytical standards											
Round	OCPs	PCB	PCDD/ PCDF	dl-PCB	PBDE	HxBB	HBCD	Toxa-phene	PFAS	PFOS precursors	Subtotal
IL2	71	76	54	57	54				37	37	**386**
IL3	128	126	66	73	70	38	34	37	41	41	**654**
IL4	99	97	56	64	45		30	23	39		**453**
Subtotal	298	299	176	194	169	38	64	60	117	78	**1 493**

Test samples of environmental and human matrices										
Round	Sediment	Fish	Human milk	Human plasma	Air (TOL)	Air (MeOH)	Water	Transfor-mer oil	Sub-total	Total Samples
IL2	75	79	89	18	64	21	32	30	408	**794**
IL3	136	105	115	26	90	19	43		534	**1 188**
IL4	110	230	124	23	80	22	33		622	**1 075**
Subtotal	321	414	328	67	234	62	108	30	1 564	**3 057**

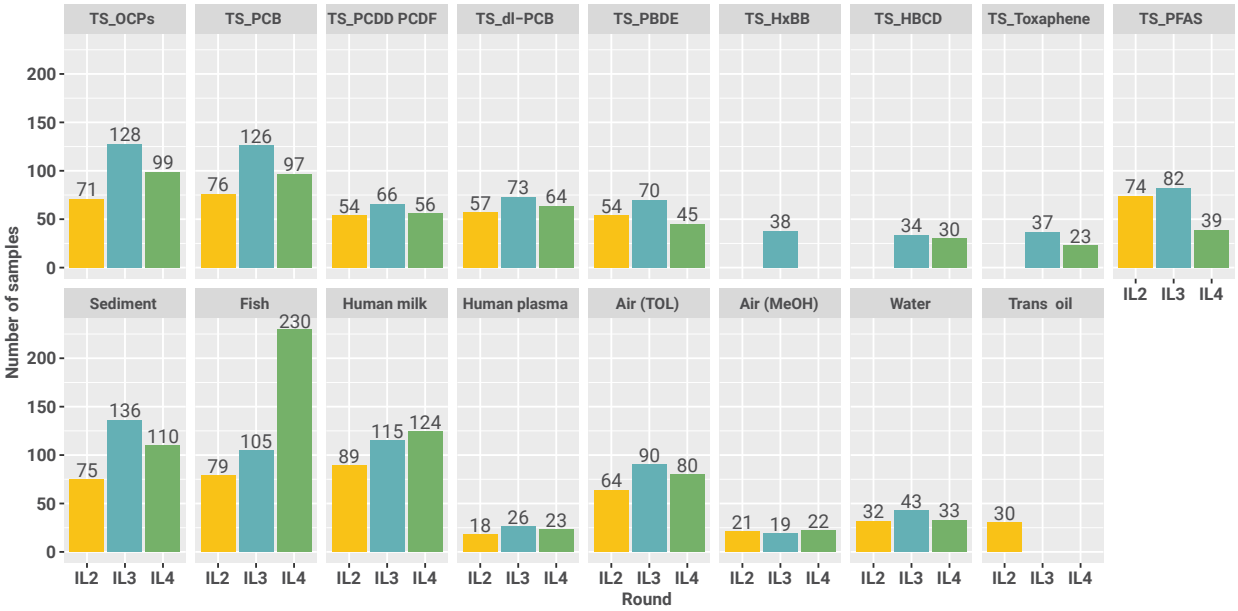


Figure 2: Number of test samples prepared and distributed per round (IL2-IL4)

3

Results: Characteristics of the participating laboratories

3.1 Identity of laboratories in the database

Our database of POP laboratories contains more than 300 laboratories that were invited for participation in the interlaboratory assessments. 289 laboratories responded on the invitation and registered at least once. Table S 3 in the Appendix provides a compilation of the laboratories in each region. The total number of countries was 82 (Figure 3). WEOG had 20 different countries in the database, followed by Africa and Asia with 18 countries each. The distribution of the 289 laboratories by country is shown in Figure 4. By far, the largest number of laboratories was found in China (42), followed by Viet Nam (16), Thailand (13), and Brazil and Colombia with 12 laboratories each.

Table S 2 summarizes the 289 laboratories that had registered for any or all of the four rounds of interlaboratory assessments (IL1-IL4), designated as "Reg_ILx" whereby x denotes the round. The four columns at right indicate, if the laboratory had submitted analytical results and obtained at least one z-score in its set of results. The columns are designated as "Res_ILx". The color codes and cells indicate laboratories that registered and obtained at least one z-score in any of the four interlaboratory assessments (Y and green highlight), laboratories that did not register (notR and yellow highlight), laboratories that registered but either not delivered results or did not obtain any z-score (N and red highlight). Blank and white cells indicate that the laboratory did not participate in the respective round.

The laboratories that registered but did never deliver results are shown in the Appendix in Table S 4; these were 61 in total. Of these, one laboratory from Africa registered 3-times and did not deliver. Seven laboratories registered twice and did never deliver, and 53 registered once and failed to deliver results. As to the regional distribution of these laboratories, 18 were from GRULAC, 16 from Asia, 14 from Africa, 7 from CEE, and 6 from WEOG.

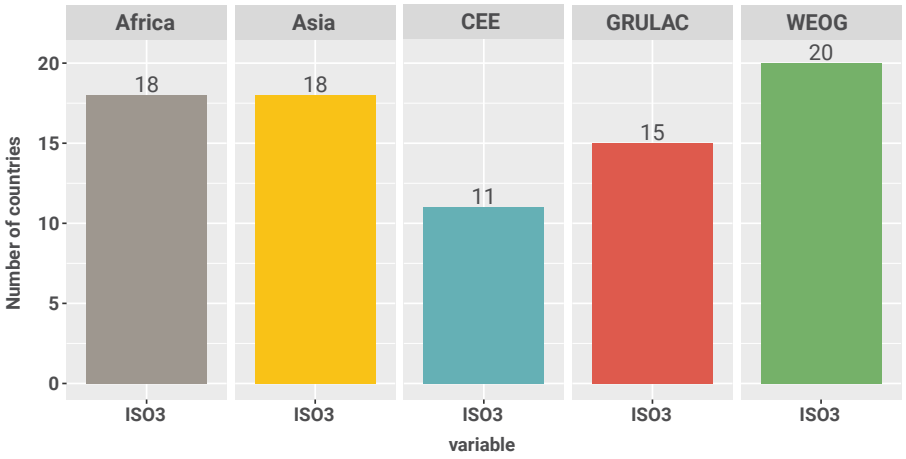


Figure 3: Number of countries within each region that had a laboratory registered (IL1-IL4)

Figure 4: Number of laboratories in each country grouped and colored by region (database of IL1-IL4)

3.2 Number of registrations vs. results

From 2010 to 2019, four rounds of interlaboratory assessments were implemented. The number of laboratories that registered and that delivered a set of results and obtained at least one z-score in the respective round are displayed in Table 5 and visualized in Figure 5. Overall, there were 532 registrations by laboratories and 420 sets of results were received as indicated in Figure 5. Thus, there were more laboratories registered than delivering results; overall, we had expected 112 more sets of results (based on registration numbers). The number of laboratories that did not deliver results was 21, 16, 43 or 32, resp. for the rounds 1-4; on average 21%.

Table 5: Overview on number of laboratories that registered vs. number of laboratories that obtained at least one z-score in each round of the interlaboratory assessments

Criterion	IL1	IL2	IL3	IL4
No of laboratories registering in the round	103	105	176	148
No of laboratories obtaining at least one z-score	82	89	133	116
No of laboratories not delivering results in the round	**21**	**16**	**43**	**32**

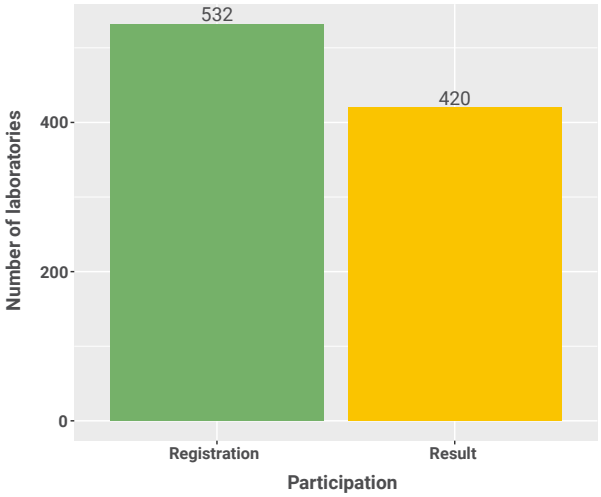


Figure 5: Total number of laboratories that registered vs. number of laboratories that delivered results in the same round

The individual participation of each laboratory as concerns registration in a specific round and delivery of results in the same specific round is provided in the Appendix as Figure S 1. Stacked red and green bars designate that the laboratory had registered in the given round and delivered a set of results. Red bars only imply that the laboratory had registered but was unable to deliver results in the same round. Since results without registration are not possible, there is no lab having green bars only. The number of stacked bars identifies the number of rounds the laboratory had participated. The maximum number of rounds to be achieved is four; this goal is achieved by laboratories having a low laboratory designation number (<L105).

3.2.1 Region as denominator for laboratory participation

Figure 6 and Table 5 details the number of registrations and number of results for each round of the interlaboratory assessment. It can be seen that IL3 had the largest number of registrations and deliveries (176 vs. 133). The number of laboratories that registered for each of the rounds was between 103 and 176; from these, between 82 and 133 sets of results were obtained in the respective round. From Table 5 and Figure 6 it can be seen that not all laboratories delivered results so that in each round about one fourth of the laboratories were too ambitious; they registered but did not manage to provide results.

A differentiation of participation with respect to registration and results by region is shown in Figure 7 and Table 6. On percentage, registration vs. results, there is not much difference between the regions.

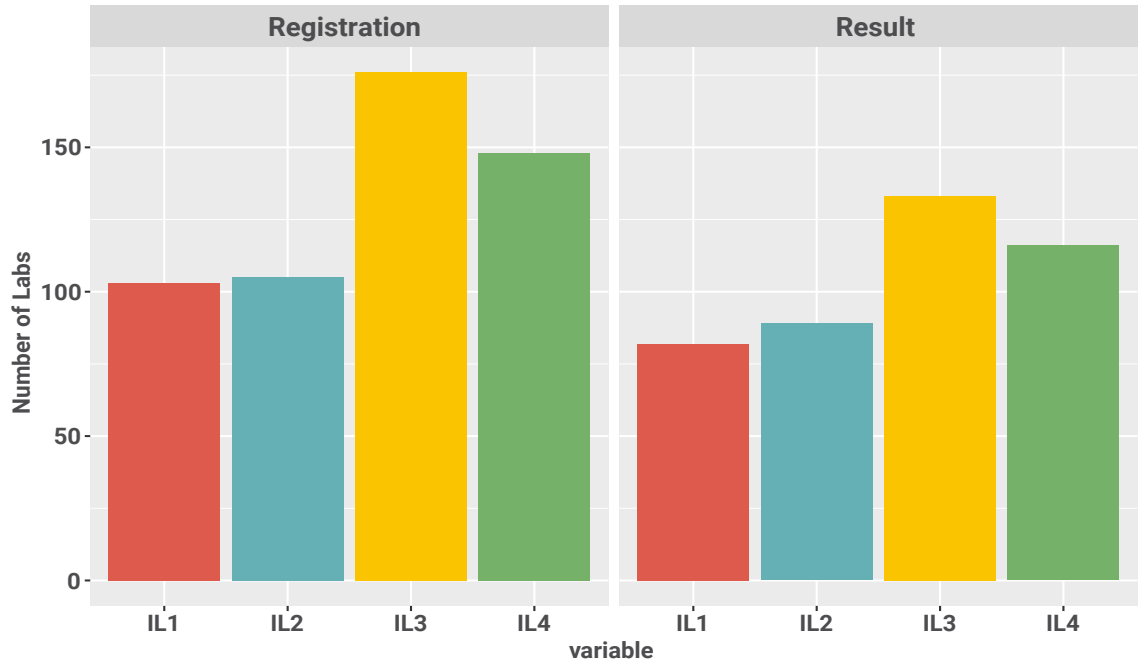


Figure 6: Number of laboratories that registered vs. number of laboratories that delivered results in each round

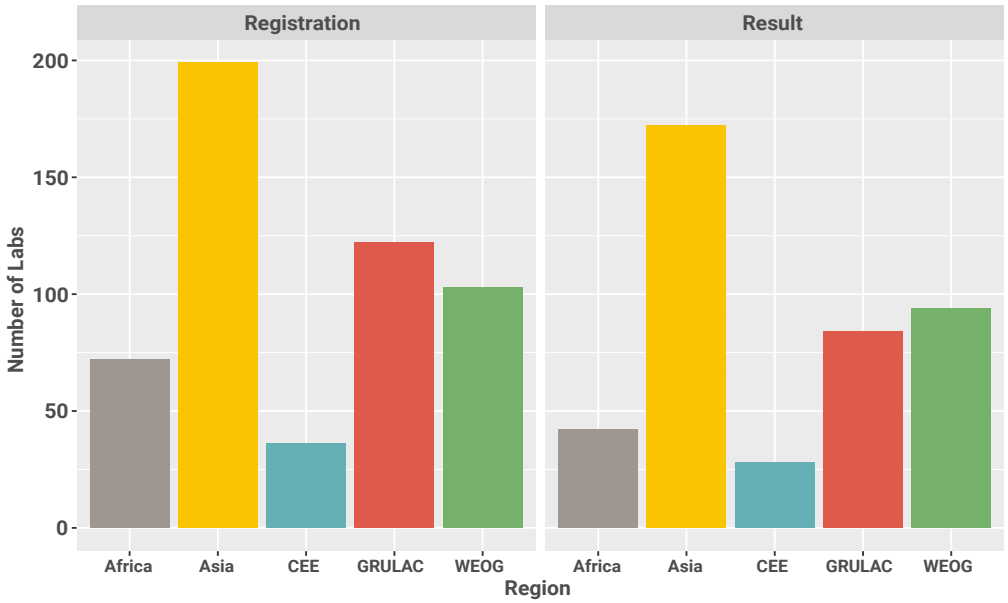


Figure 7: Overview on number of laboratories that registered vs. number of laboratories that delivered results by region

Table 7 groups the 532 registrations from laboratories and 420 sets of results delivered into the respective region where the laboratory was located. IL3 had the highest participation with 176 laboratories registered but only 133 delivered results; thus, only 76% delivered and almost one quarter of the laboratories failed. IL1 had 103 laboratories registered and 82 sets of results (80% delivered), for IL2, the ratio was 105 to 89 or 78%, and IL4 had 148 registrations toward 116 deliveries, corresponding to 78%. Almost no improvement was made from IL1 to IL4, so that in general about 20% (or more) of the efforts failed.

Of the 289 laboratories registered, 228 laboratories from 72 countries delivered results and obtained at least one z-score in any of the four rounds (IL1-IL4). In terms of registration and delivery of results, most laboratories were from Asia-Pacific region with 199 registrations of laboratories and 172 laboratories that had obtained at least one z-score. The percentage of laboratories that registered but did not succeed to obtain any z-score may be viewed as being too ambitious or overestimating the own capacity. In total, 14% of the Asian laboratories did not succeed in the interlab assessment. The best achievement was for WEOG laboratories, where the delivery rate was 91% and only 9% of the registered laboratories failed to deliver results. The worst performance was obtained for the African laboratories where 42% did not succeed (72-times registrations and only 42 sets of results delivered by the laboratories). Overall, 21% or 61 laboratories registered but did not deliver. These are shown in Table 6. The numbers for the laboratories are visualized according to the regions in Figure 8.

Table 6: Number of registrations by laboratories vs. number of laboratories that obtained at least one z-score in each round of the interlaboratory assessments

Row Labels	Number of Labs Registered	Number of Labs with Results	Percentage with no Results
Africa	72	42	42%
Asia	199	172	14%
CEE	36	28	22%
GRULAC	122	84	31%
WEOG	103	94	9%
Total	**532**	**420**	**21%**

Table 7: Summary of number of laboratories registered and delivering results according to UN region and round

Round	IL1		IL2		IL3		IL4		Total	
Region	Reg	Result	Reg	Result	Reg	Result	Reg	Result	Reg	Result
Africa	17	10	12	5	19	14	24	13	72	42
Asia	38	33	45	42	68	53	48	44	199	172
CEE	3	3	4	4	23	16	6	5	36	28
GRULAC	32	23	14	11	39	25	37	25	122	84
WEOG	13	13	30	27	27	25	33	29	103	94
Total	103	82	105	89	176	133	148	116	532	420

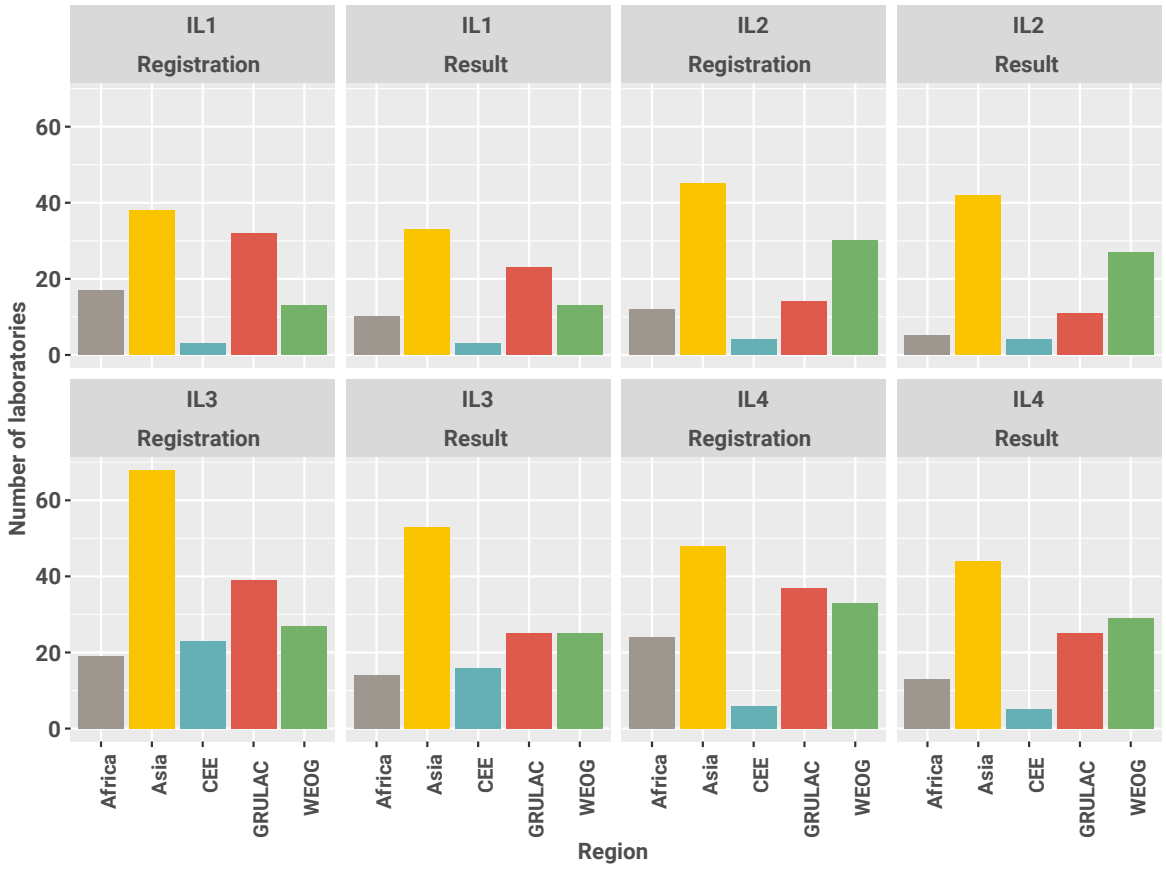


Figure 8: Number of laboratories registered and delivering results for each round and by region

3.2.2 Country as denominator for laboratory participation

Table 8 groups the laboratories according to their corresponding regions. In total, the 532 registrations were from laboratories in 82 countries and the 420 sets of results came from laboratories in 72 countries. Across the four rounds, 20 countries were from WEOG and 18 each from Africa and Asia. Since the identity of the laboratories is not disclosed, laboratories, if they wish to do so, must identify themselves to their or other countries or UNEP.

Table 8: Summary of number of countries with laboratories registered and delivering results according to UN region and round

Round	IL1		IL2		IL3		IL4		Total	
Region	Reg	Result	Reg	Result	Reg	Result	Reg	Result	Reg	Result
Africa	12	7	10	5	13	9	14	10	18	12
Asia	11	8	10	9	13	13	13	13	18	17
CEE	2	2	2	2	10	8	5	5	11	9
GRULAC	12	11	10	8	14	12	15	12	15	14
WEOG	10	10	16	15	15	14	16	15	20	20
Total	47	38	48	39	65	56	63	55	82	72

The 289 laboratories were from 82 countries as shown in Figure 9. It can be seen that laboratories from ten countries – Albania, Barbados, Cameroon, Democratic Republic of the Congo, Ethiopia, Georgia, Madagascar, Malaysia, Sudan, and United Republic of Tanzania - did never achieve to obtain z-scores; thus, 72 countries had laboratories that delivered results and obtained z-sores.

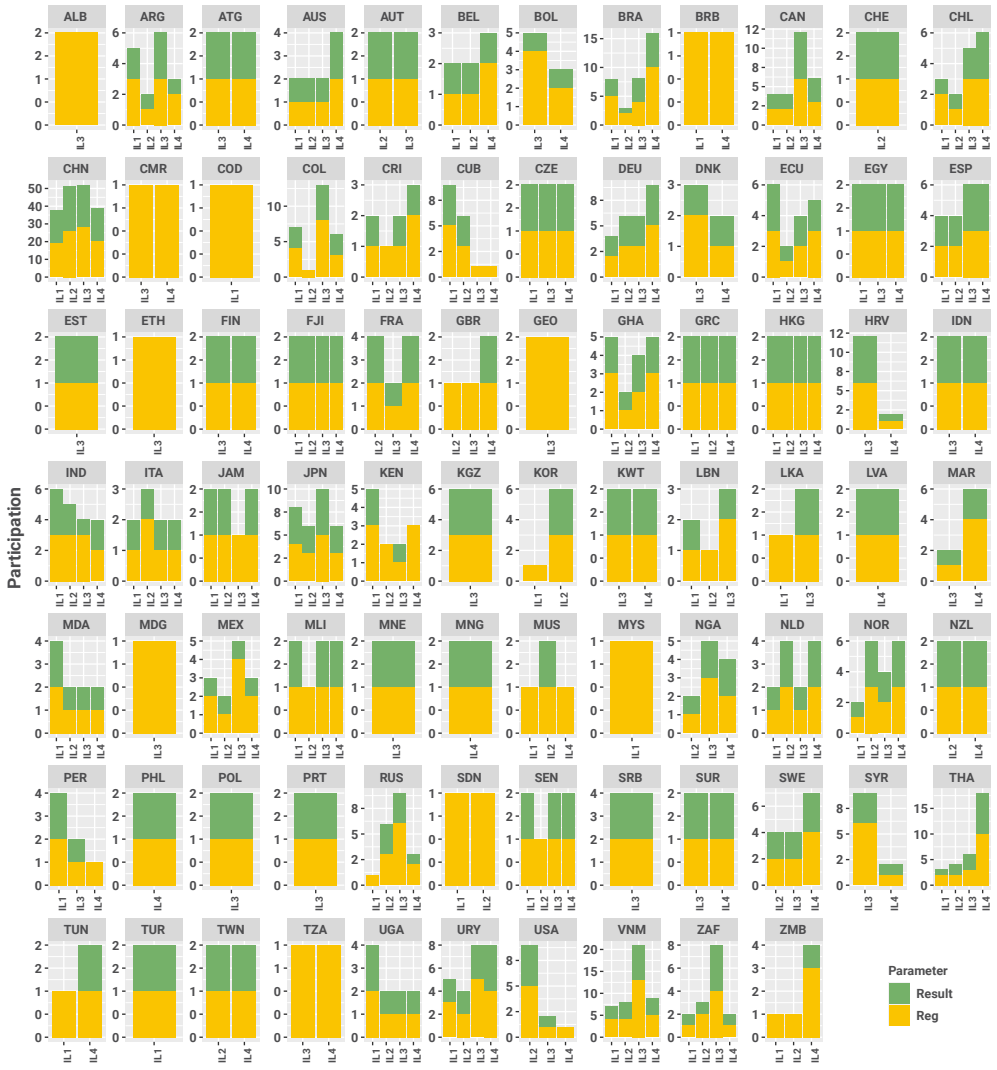


Figure 9: Graphical sketch showing the number of laboratories registered and delivering results in each country

The 228 laboratories that obtained z-scores were in 72 different countries whereas participation was from 82 countries. The ten countries where the laboratories did not deliver can be identified by the missing bar in the facets containing "Result". Further the round where the participation took place is shown by the respective color of the bars. Overall, most countries were from WEOG, followed by Asia. The country participation and results delivery by all laboratories in the given country is shown in Figure 10.

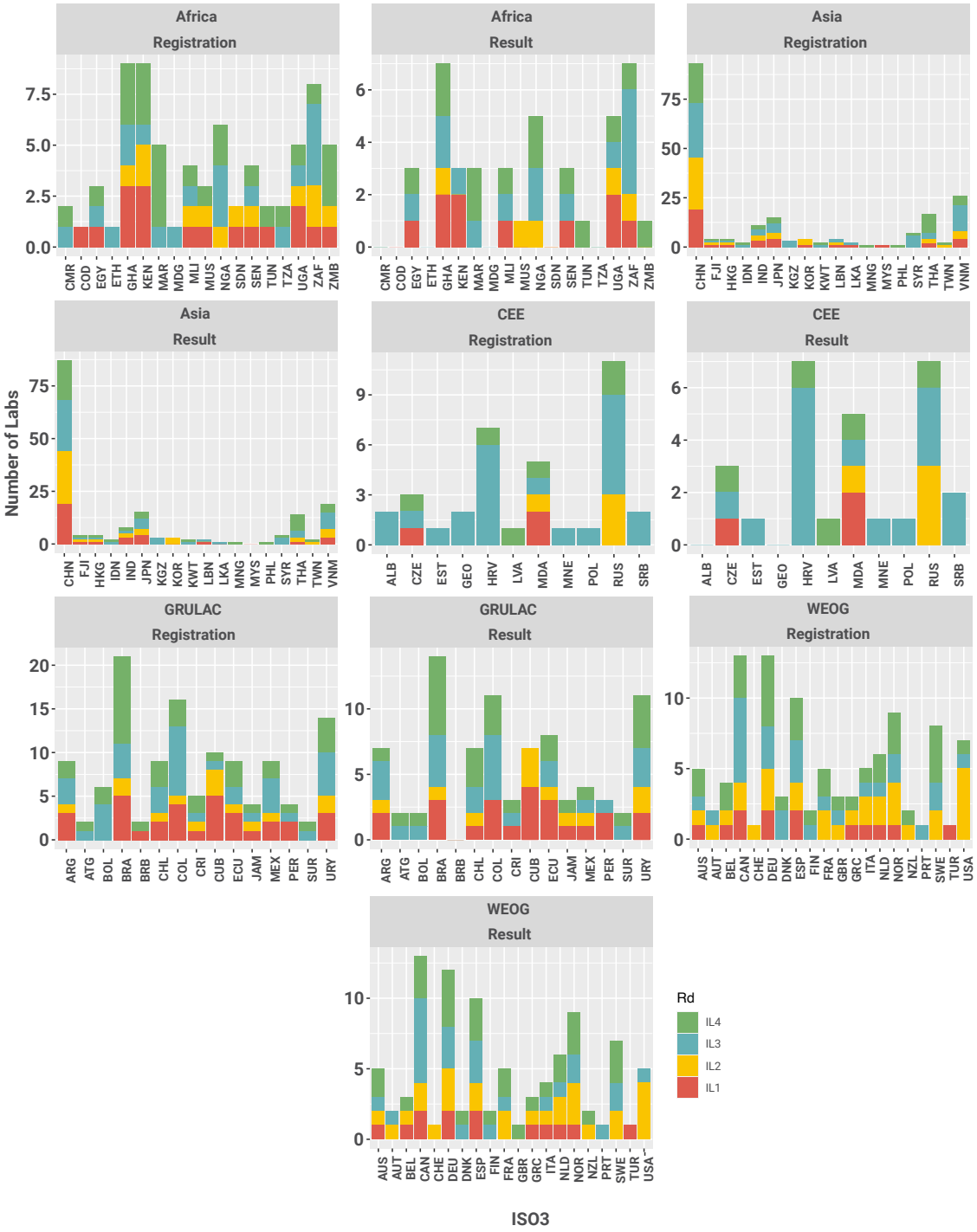


Figure 10: Laboratories *per* country showing registrations and results grouped into region and colored by round

4

Performance as z-scores

4.1 Summary of z-scores

In summary, there were 41 575 z-scores assigned to the 228 laboratories that delivered results. The z-scores are grouped into the UN regions and shown for the four rounds (Table 9). The z-scores were distributed as shown in Figure 11. It can be seen that the largest number of z-scores was always obtained in IL3. Among the POPs groups, the dl-POPs had the largest shares in Asia, CEE, and WEOG whereas the OCPs had the largest shares in Africa and GRULAC.

Table 9: Number of z-scores by round (IL1-IL4) and region

	Africa (N=2295)	Asia (N=19144)	CEE (N=2430)	GRULAC (N=4568)	WEOG (N=13138)	Overall (N=41575)
			Round			
IL1	394	3 522	154	948	1 446	6 464
IL2	191	5 149	512	695	3 944	10 491
IL3	1 054	5 884	1 401	1 338	3 578	13 255
IL4	656	4 589	363	1 587	4 170	11 365

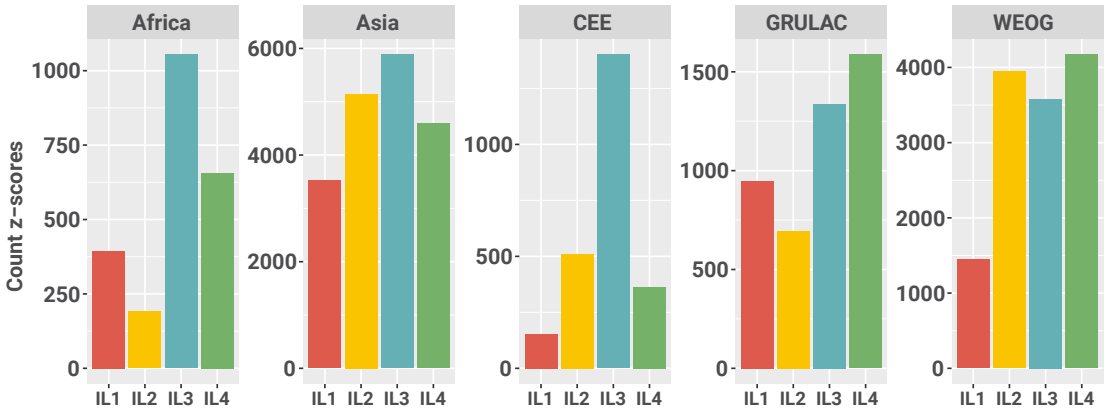


Figure 11: Number of z-scores obtained by laboratories in each round and grouped by UN region

In total, 41 575 z-scores were assigned in the four rounds (Table 10). It can be seen that not all the POPs groups were tested from the beginning and that some matrices, such as transformer oil (TO in IL2) or fly ash (Ash as a proxy for the air extract in IL1) were included only once. The numbers of z-scores were largest in IL3, due to the largest number of participating laboratories. The test solutions of the analytical standards (Test solution) always had the highest number of results (>15 000). Among the POPs_groups, most z-scores were attributed to the dl-POPs (19 500), followed by the OCPs (9 526).

Table 10: Summary of z-scores by round, type of test sample, matrix, and POP-group

	IL1 (N=6 464)	IL2 (N=10 491)	IL3 (N=13 255)	IL4 (N=11 365)	Overall (N=41 575)
Type					
Test solution	2 434	3 542	4 834	4 202	15 012
Abiotic	2 175	4 113	4 847	4 822	15 957
Biota	1 855	2 836	3 574	2 341	10 606
Matrix					
Test solution	2 434	3 542	4 834	4 202	15 012
Air	0	1 974	2 481	2 450	6 905
Sediment	1 315	1 975	2 325	2 120	7 735
Fish	955	1 360	2 176	866	5 357
Human milk	900	1 404	1 295	1 313	4 912
Water		20	41	252	313
Human plasma		72	103	162	337
Ash	860				860
Transformer oil		144			144
POP_group					
OCPs	1 599	2 160	3 377	2 390	9 526
indPOP	966	1 716	2 202	1 497	6 381
dl_POPs	3 899	5 121	5 897	4 613	19 530
BFR		1 074	1 149	996	3 219
PFAS		420	630	1 869	2 919

4.2 Performance assessment using z-scores

In the four rounds of interlaboratory assessments, 41 575 z-scores have been attributed to laboratories for their performance in the POPs analysis. The following sections, tables and figures provide some details on this abundance of data.

4.2.1 General overview on quality of the z-scores

Across all ILs, 41 575 z-scores were assigned. The distribution of the z-scores according to their quality, corresponding to performance of the laboratory, results in: 25 192 z-scores for S, 3 991 for Q, 10 305 for U, and 584 and 1 503 for C and I, respectively.

4.2.2 Regional performance as to quality of z-scores

The regional distribution and the quality of the 41 575 z-scores is shown in Table 11 and in Figure 12. 46% or 19 144 of all z-scores are from laboratories located in the Asian region; 32% (N=13 138) from WEOG, 11% (N=4 568) from GRULAC, and only 6% to each, Africa (N=2 295) and CEE (N=2 430). The laboratories in the Asian region also generated 49% of all the S results; WEOG had 36%. All other regions are negligible as to the satisfactory results.

It shall be noted that the two successful regions – Asia and WEOG – also generated the highest percentage of the unsuccessful results: Among the U results, Asia has 43%, WEOG has 20%, GRULAC has 16%, Africa 13%, and CEE 7.7%. The distribution of the z-scores by UN region is shown in Figure 12.

Table 11: Quality of the z-scores by region (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

z-score	Africa (N=2 295)	Asia (N=19 144)	CEE (N=2 430)	GRULAC (N=4 568)	WEOG (N=13 138)	Overall (N=41 575)
S	480	12 314	1 249	2 053	9 096	25 192
Q	199	1 744	224	431	1 393	3 991
U	1 351	4 424	788	1 693	2 049	10 305
C	18	175	16	87	288	584
I	247	487	153	304	312	1 503

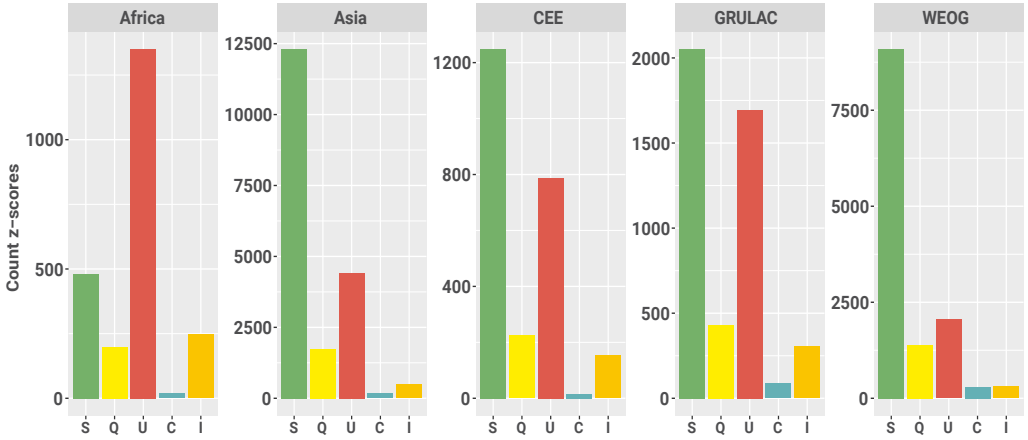


Figure 12: Number and quality of z-scores obtained by laboratories grouped by UN region

4.2.3 Time trends as to performance (quality of z-scores)

For time trends, the following observations can be drawn (Table 12 and Figure 13):

i. The number of results generated increased strongly (steeper than the number of participating laboratories) and peaked in IL3 with 13,255 assigned z-scores in total;

ii. The number of satisfactory results that were generated in subsequent rounds has increased since the inception of the interlaboratory assessments;

iii. However, expressed as percentage of total results, the overall picture is somewhat disappointing and shows opposite trends: whereas in IL1 the percentage of satisfactory results was 68%, the percentage decreased constantly to 56% in IL4, the percentage of unsatisfactory results increased from 21% in IL1 to 27% in IL4.

Table 12: Number of z-scores *per* round (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

	IL1 (N=6464)	IL2 (N=10491)	IL3 (N=13255)	IL4 (N=11365)	Overall (N=41575)
S	4 410	6 708	7 737	6 337	25 192
Q	666	1 057	1 207	1 061	3 991
U	1 388	2 237	3 570	3 110	10 305
C		153	128	303	584
I		336	613	554	1 503

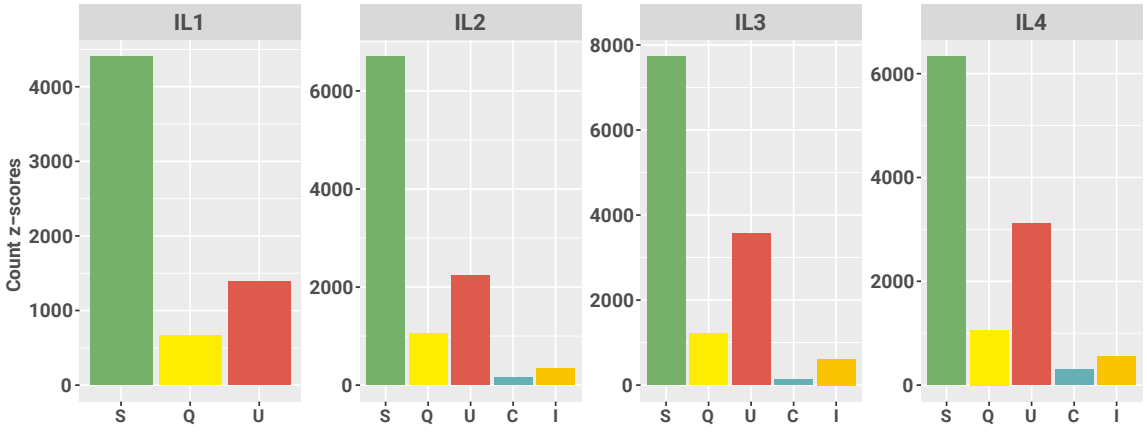


Figure 13: Number of z-scores across four rounds of interlaboratory assessments (IL1-IL4) shown as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent

The capacity for POPs analysis by POP and region are shown in Table 13 for organochlorine pesticides (OCPs), Table 14 for industrial POPs (indPOPs), dl-POPs, and PFAS listed in the Stockholm Convention, and Table 15 for the sum parameters and the overall number of z-scores.

Table 13: Number of z-scores for OCPs by POP and UN region

	aldrin	dieldrin	endrin	chlordane	DDT	heptachlor	mirex	toxaphene	a_HCH	b_HCH	lindane	endosulfan
Africa	29	49	28	106	337	89	18	0	36	64	34	66
Asia	121	141	94	601	1 335	380	117	43	115	148	97	155
CEE	20	25	16	68	239	49	14	5	28	32	26	40
GRULAC	74	87	55	282	723	182	63	26	70	90	70	144
WEOG	55	79	45	391	760	180	84	41	79	95	59	105
Subtotal	**299**	**381**	**238**	**1 448**	**3 394**	**880**	**296**	**115**	**328**	**429**	**286**	**510**

Table 14: Number of z-scores for industrial POPs, dl-POPs, and POP-PFAS by POP and UN region

	PCB6	HCB	PeCBz	PCDD	PCDF	dl_PCB	PBDE	PBDE_209	HBCD	HxBB	PFOS	PFOA	PFHxS
Africa	555	37	6	101	132	232	81	0	16	11	9	1	2
Asia	2 105	176	67	2 336	3 146	3 496	1 144	25	104	52	281	59	55
CEE	519	47	16	169	227	430	125	3	0	1	31	9	3
GRULAC	960	104	28	231	321	407	144	6	0	8	14	3	3
WEOG	1 565	127	69	1 232	1 659	2 182	1 011	36	113	53	488	96	102
Subtotal	**5 704**	**491**	**186**	**4 069**	**5 485**	**6 747**	**2 505**	**70**	**233**	**125**	**823**	**168**	**165**

Table 15: Number of z-scores for sum parameters and total by POP and UN region

	sum_drins	sum_HCH	TEQ_DF	TEQ_PCB	TEQ_total	sum_PBDE	PFOSprec	nonSC_PFAS	sum_PFOSprec	sum_nonSC_PFAS	Overall
Africa	63	64	28	40	34	12	1	12	0	2	**2 295**
Asia	193	127	654	566	540	120	58	429	11	53	**19 144**
CEE	37	49	50	58	48	14	0	23	0	9	**2 430**
GRULAC	114	100	66	63	58	20	11	34	3	4	**4 568**
WEOG	85	90	351	342	331	120	140	825	29	119	**13 138**
Subtotal	**492**	**430**	**1 149**	**1 069**	**1 011**	**286**	**210**	**1 323**	**43**	**187**	**41 575**

4.2.4 Performance assessment by POP

An overview on the z-scores for each POP is shown in Table 16. It can be seen that for chlordecone and HCBD, no z-scores could be assigned. Overall, 25 192 of the 41 575 z-scores were satisfactory; corresponding to 61%. Unsatisfactory were 25%, which means that one quarter of all results submitted by the laboratories had a coefficient of variation of greater than 2x37.5% or 75% off the agreed value (AV). The performance by POP is shown graphically in Figure 14 with the percentage according to the quality of the z-scores. POPs with a large share of green color in the stacked bars designate good performances (S) for the results and red bars, poor performances. The most favourable ratio was obtained for HBCD (80% were S). Also on the very positive side were the dl-POPs and especially the TEQs but also toxaphene and the PFAS. Among the dl-POPs, the performance for dl-PCB was lower than for PCDD and PCDF. For toxaphene, HBCD, PFOA, PFHxS, sum_PBDE and the PFAS sum parameters, it must be noted that the overall number of z-scores was much lower than for the dl-POPs and their sum parameters.

From Figure 14 it can also be seen that for some POPs, there were more unsatisfactory (U) results and satisfactory (S) results: these were endosulfan (S/U ratio = 0.78), sum drins (S/U ratio = 0.97), and sum HCHs (S/U ratio = 0.80); for were β-HCH, the ratio S/U was 1.03 (just positive).

Figure 15 shows the number of z-scores colored according to the POP group. It can be seen that most z-scores were obtained for dl-POPs, followed by OCPs. Figure 16 shows the z-scores as isolated bar graphs for each POP.

For OCPs and PCB6, the performance is disappointing: roughly only half of the OCP (46%) and industrial chlorinated POPs (indPOPs) (51%) results are satisfactory. For dl-POPs, the percentage was 69% for BFRs 63%, and for PFAS was 70%.

Table 16: Overview on z-sores across four rounds of interlaboratory assessments (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

	S (N=25 192)	Q (N=3 991)	U (N=10 305)	C (N=584)	I (N=1 503)
Organochlorine pesticides (OCPs)					
aldrin	148	31	89	9	22
dieldrin	168	39	127	10	37
endrin	101	35	84	2	16
chlordane	776	116	381	64	111
DDT	1 533	337	1 203	57	264
heptachlor	421	81	306	14	58
mirex	169	25	82	4	16
toxaphene	91	8	16	0	0
α-HCH	141	32	116	5	34
β-HCH	152	36	148	18	75
lindane	136	17	96	3	34
chlordecone					
endosulfan	179	75	234	0	22
Industrial chlorinated POPs (indPOPs)					
PCB6	2 923	641	1 907	36	197
HCB	240	52	162	8	29
PeCBz	100	11	47	11	17
HCBD					
Dioxin-like POPs (dl-POPs)					
PCDD	2 900	402	612	67	88
PCDF	3 948	489	865	93	90
dl-PCB	4 250	631	1 611	48	207
Brominated flame retardants (BFRs)					
PBDE	1 531	247	604	44	79
PBDE_209	32	4	21	6	7
HBCD	186	7	32	4	4
HxBB	82	11	17	7	8
Perfluorinated alkyl substances (PFAS)					
PFOS	569	61	168	16	9
PFOA	122	16	22	4	4
PFHxS	116	11	18	8	12
Sum parameters (calculated from the above POPs)					
sum_drins	215	56	221		
sum_HCHs	176	35	219		
TEQ_DF	889	105	155		
TEQ_PCB	702	105	262		
TEQ_total	755	84	172		
sum_PBDE	199	23	64		
PFOSprec	150	19	38	2	1
nonSC_PFAS	935	122	160	44	62
sum_PFOSprec	28	7	8		
sum_nonSC_PFAS	129	20	38		

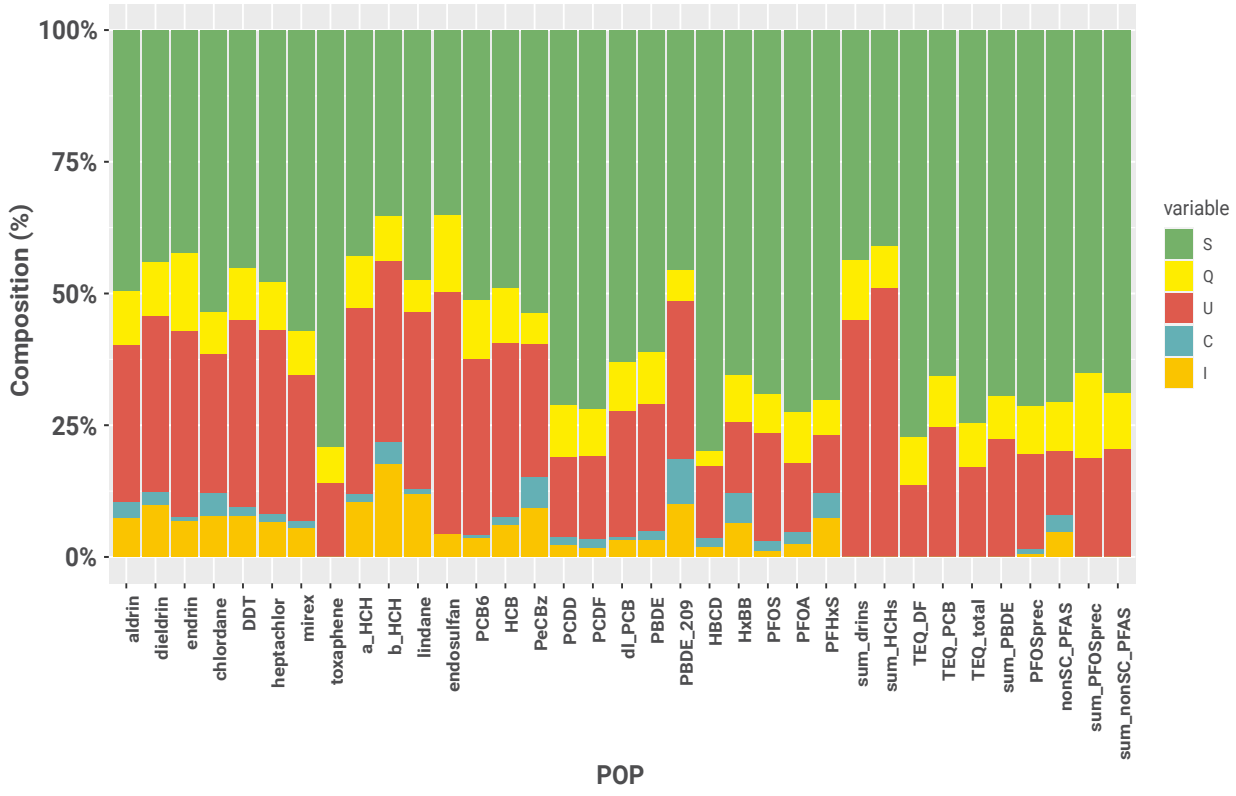


Figure 14: Performance (in percent) of laboratories by POP and z-score (IL1-IL4)

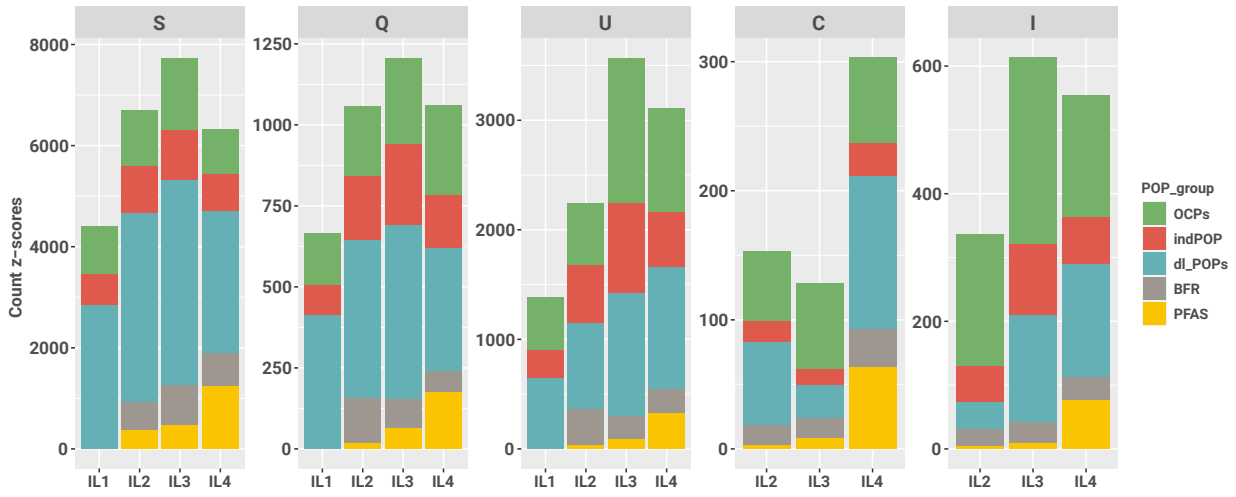


Figure 15: Number of z-scores (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent) by score and round (IL1-IL4)

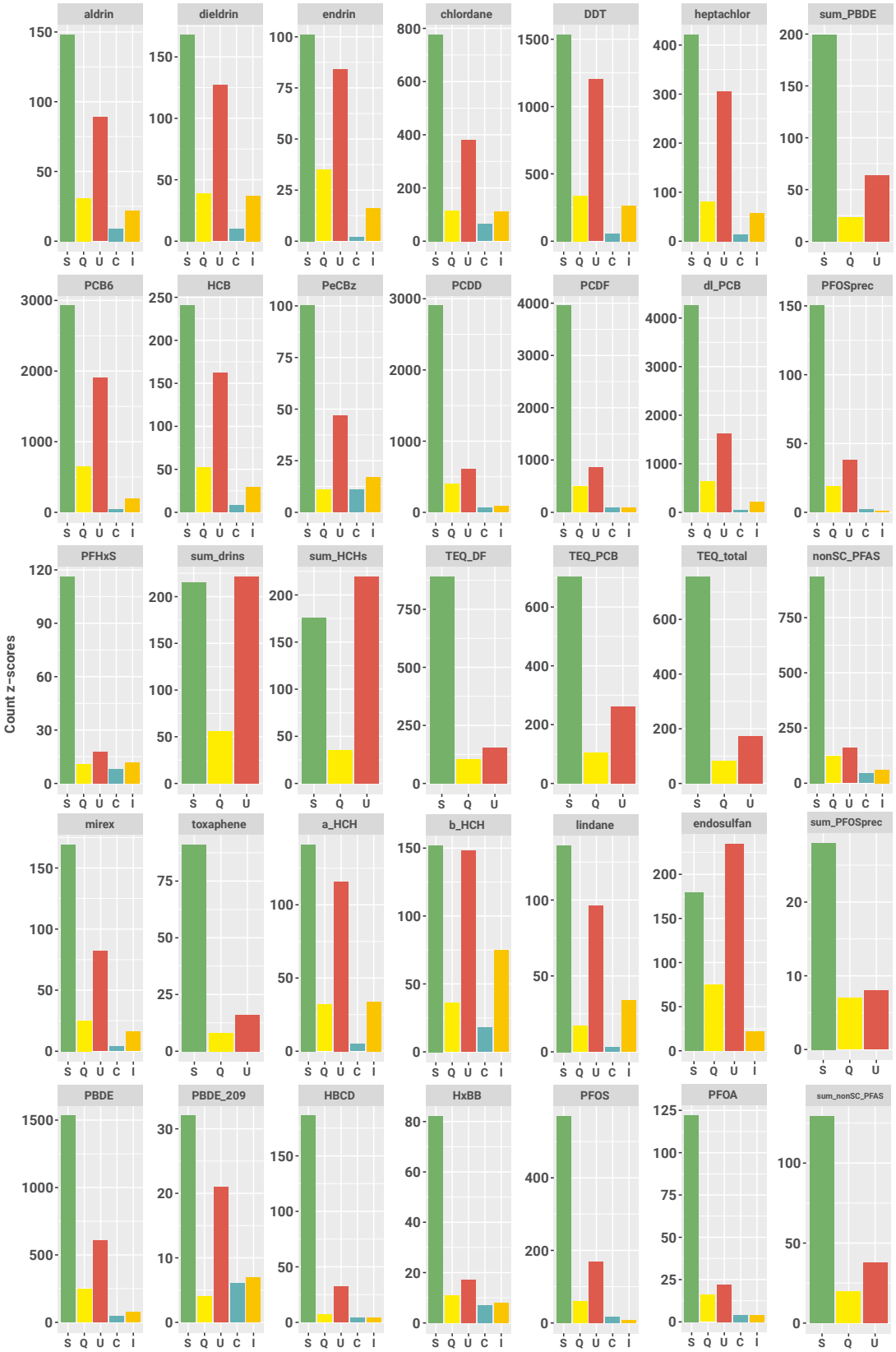


Figure 16: Performance of laboratories by POP and z-score (IL1-IL4) (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

4.2.5 Performance assessment according to test type and matrix

Performance by type of test sample

As mentioned in section 2.1, there were two broad kinds of test materials: test solutions of analytical standard and naturally contaminated test samples; the latter ones can be divided into abiotic and biota samples. The matrices included into these categories are shown in Table 1 and Figure 2. Abiotic test samples included air extract, water, fly ash, and transformer oil; biota test samples included human milk, fish, and human plasma.

Table 17 summarizes the number and quality of the z-scores according to type of test sample. Abiotic samples generated slightly more z-scores than test solutions; biota had much less. Accordingly, the test solutions generated the best results with 67% assigned satisfactory followed by abiotic samples (59%), and biota (53%), see Figure 17.

The quality of the z-scores by type and colored by region is shown in Figure 18. A table further disaggregating the number of z-scores according to the types of test samples for each POP is contained in the Appendix as Table S 5.

Table 17: Summary of z-sores across four rounds of interlaboratory assessments by type (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

	S	Q	U	C	I	Subtotal
Test solution	10 126	1 550	3 173	16	147	**15 012**
Abiotic	9 413	1 466	4 165	193	720	**15 957**
Biota	5 653	975	2 967	375	636	**10 606**
Overall	**25 192**	**3 991**	**10 305**	**584**	**1 503**	**41 575**

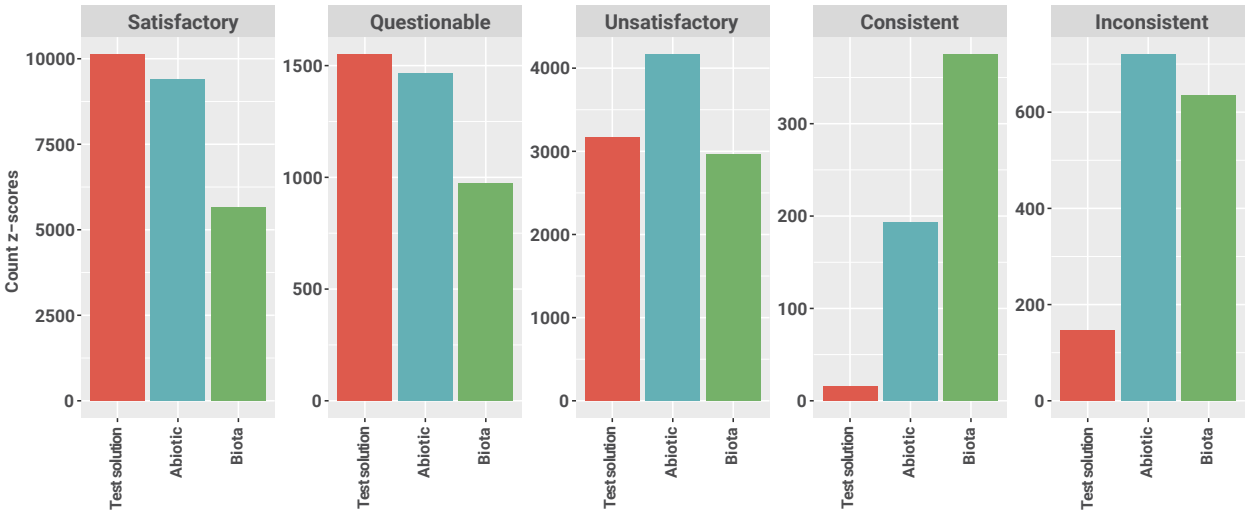


Figure 17: Quality of z-scores according to type of sample (IL1-IL4)

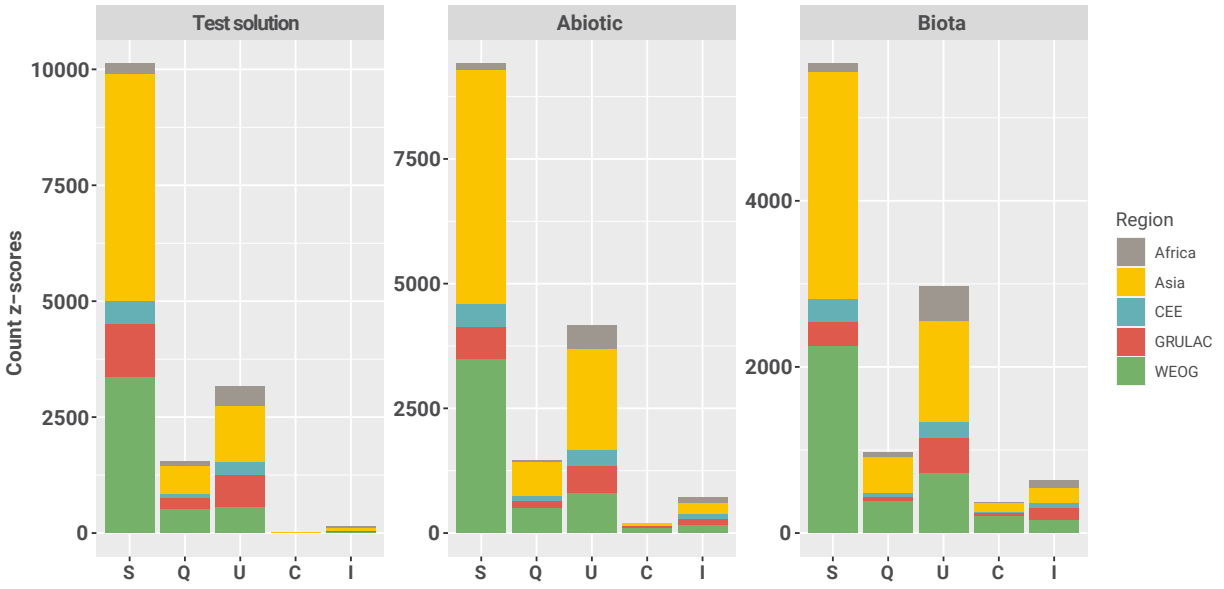


Figure 18: Performance of laboratories by POP and type of test sample, stacked bars (IL1-IL4) (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

Performance by matrix of test sample

The quality of z-scores for each POP and by matrix is shown in Table 18. Accordingly, the human plasma matrix was assigned most of the satisfactory results (73%), followed by the test solutions (67%). Poorest results were obtained for fish with a ratio of 1.67 comparing S/U (2 740/2 194) followed by water (1.87) and sediment (1.97). For details see Figure 19.

The quality of the z-scores by type and colored by region is shown in Figure 20. Whereas typically, Asia and WEOG have same capacities and performance, it can be seen that in WEOG, there is more capacity and better performance for water (and PFAS).

Figure 21 provides further insight by showing the performance of the laboratories by test matrix and POP.

Table 18: Summary of z-sores across four rounds of interlaboratory assessments according to matrix of the test sample

	Satisfactory	Questionable	Unsatisfactory	Consistent	Inconsistent	Subtotal
Test solution	10 126	1 550	3 173	16	147	**15 012**
Sediment	4 333	732	2 194	107	369	**7 735**
Water	173	28	93	9	10	**313**
Air	4 300	567	1 620	77	341	**6 905**
Ash	545	117	198			**860**
Transformer oil	62	22	60			**144**
Fish	2 740	504	1 641	206	266	**5 357**
Human milk	2 666	442	1 292	160	352	**4 912**
Human plasma	247	29	34	9	18	**337**
Overall	**25 192**	**3 991**	**10 305**	**584**	**1 503**	**41 575**

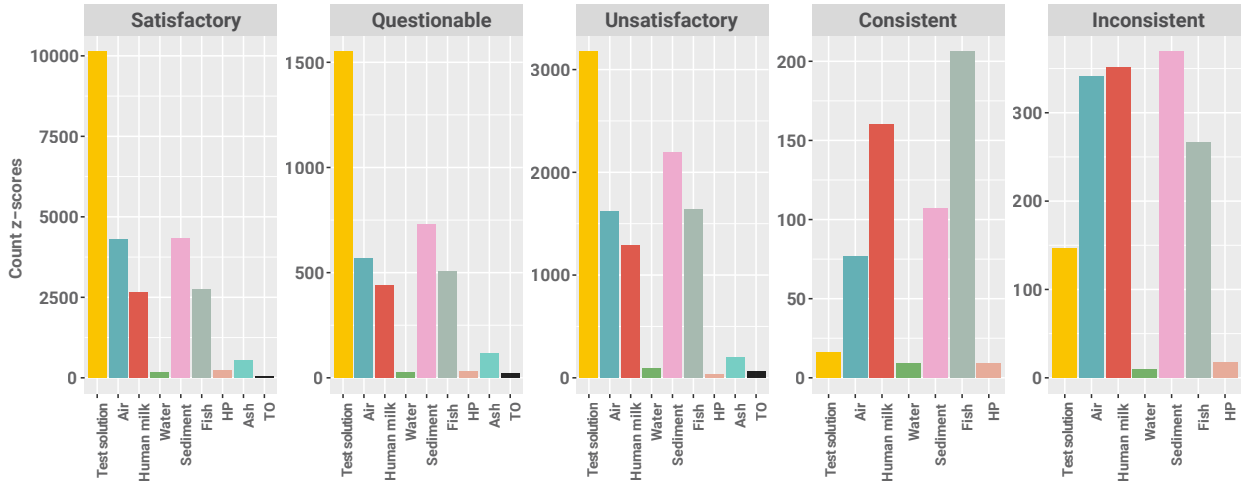


Figure 19: Quality of z-scores according to matrix of sample (IL1-IL4)

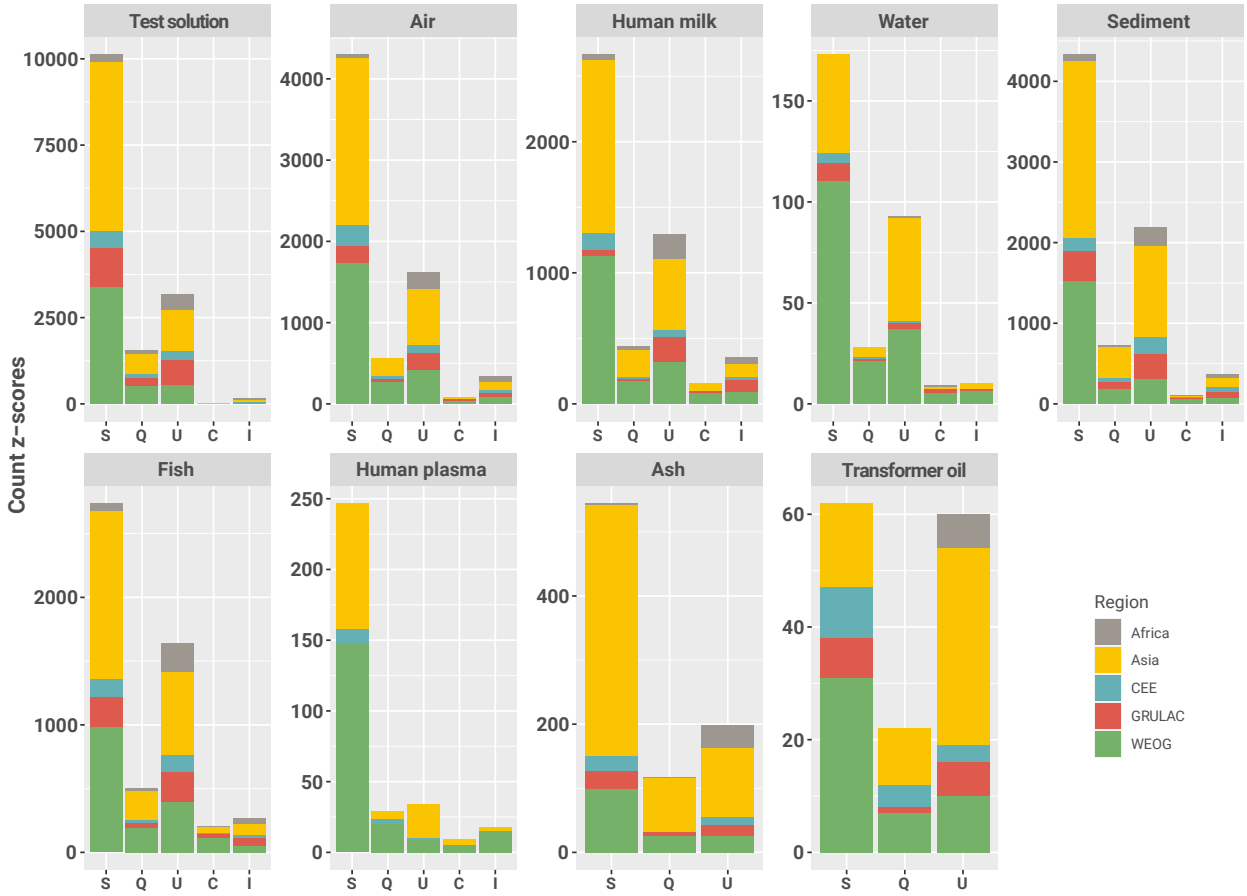


Figure 20: Performance of laboratories by matrix and colored by region, stacked bars (IL1-IL4) (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

Figure 21: Performance of laboratories by POP and test matrix, stacked bars (IL1-IL4) (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

4.2.6 Performance by country

The following Figure 22 displays the POPs analytical capacity by country and the participation in the interlaboratory assessments. Shown are the 72 countries where the laboratories delivered results. It can be seen that accordingly some countries did not have capacity for POPs analysis for any of the types of test samples, such as Antigua and Barbuda, Ghana, Kyrgyzstan, Mauritius, Nigeria, or Zambia. Other, also developed countries, such as before all China, but also Brazil, have very good capacity. In countries, like Argentina, Chile, Ecuador, Egypt, India, Kuwait, Russian Federation, Thailand, Uruguay, Viet Nam, and South Africa, capacity is being build and is strongest for the test solutions and abiotic samples.

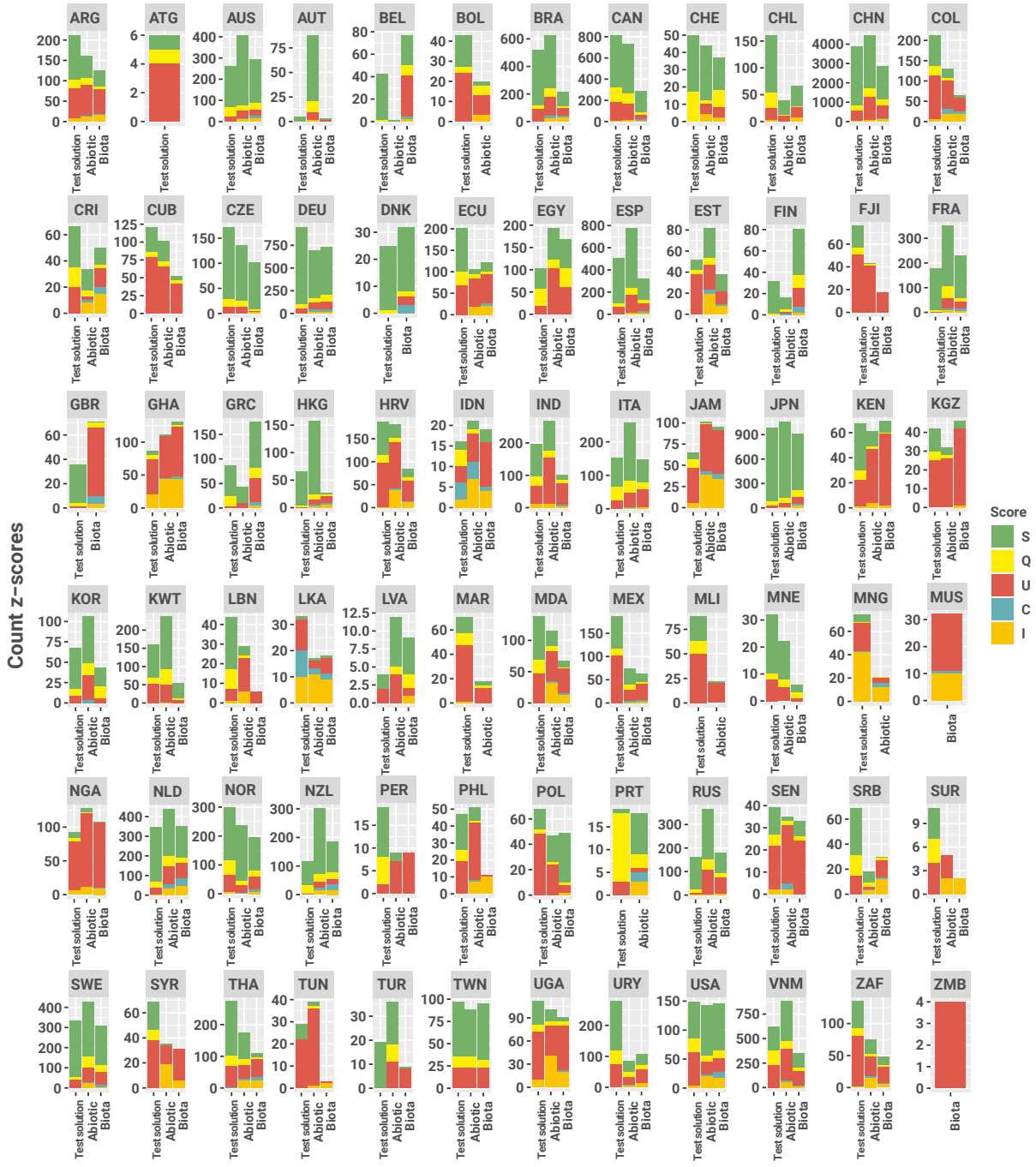


Figure 22: Performance of laboratories by country and test sample type (IL1-IL4) (as z-score interpretation with S=satisfactory, Q=Questionable, U=Unsatisfactory, C=Consistent, I=Inconsistent)

5

Trends across the four rounds

Across the four rounds of interlaboratory assessment by using the number of participations of a laboratory as an indicator, we made two assessments: (i) on the commitment of the laboratories and (b) their "learning" performance can be made.

5.1 Reliability by laboratory's participation as indicator

Among the 289 laboratories, 61 did never deliver results, 119 delivered once, 52 twice, 31 laboratories provided three-datasets of results, and 26 laboratories registered for all four rounds and delivered 4-times. Their performance is further assessed in section 5.3.

5.2 Registrations and results as indicators

The summary for laboratories according to the number of their participation is discussed first. Table 19 shows that among the 289 laboratories that registered in either round, there was an overall difference of 122 registrations by laboratories that received the test materials but did not succeed to deliver results. This accumulated failure corresponds to 42%. Among the 4-times or 3-times participating laboratories, there were only eight or six laboratories that did not deliver results; corresponding to 3% and 2% failure, resp. Less experienced laboratories had higher quota; 5% for laboratories that registered twice and 11% for laboratories that registered once and then did not deliver results. Overall, there were 61 laboratories that registered and did not deliver a result in any of the rounds.

Table 19: Multiple registrations by the same laboratory: Summary of registrations vs. delivery of results

Number of Participations	Results	Registration	Difference
4x	26	34	8
3x	31	37	6
2x	52	67	15
1x	119	151	32
0x	0	61	61
Total	289	289	122

From the 289 laboratories, there were 26 and 31 laboratories that delivered results 4-times or 3-times (Table 20), 52 laboratories twice, 119 laboratories once, and 61 registered but did not deliver results.

These do not show up in the z-scores. The summary of laboratories that delivered results are summarized in Table 21 by grouping into the UN regions.

Table 20: Multiple registration: Number of laboratories by region that registered

Number of Participations	Africa	Asia	CEE	GRULAC	WEOG	Sub-total
1x	16	57	22	27	29	151
2x	12	18	2	21	14	67
3x	4	14	2	7	10	37
4x	5	16	1	8	4	34
Total	**37**	**105**	**27**	**63**	**57**	**289**

Table 21: Multiple registration: Number of laboratories by region that delivered results/obtained z-scores

Number of Participations	Africa	Asia	CEE	GRULAC	WEOG	Sub-total
0x	14	16	7	18	6	61
1x	11	46	16	21	25	119
2x	6	18	1	14	13	52
3x	5	10	2	5	9	31
4x	1	15	1	5	4	26
Total	**37**	**105**	**27**	**63**	**57**	**289**

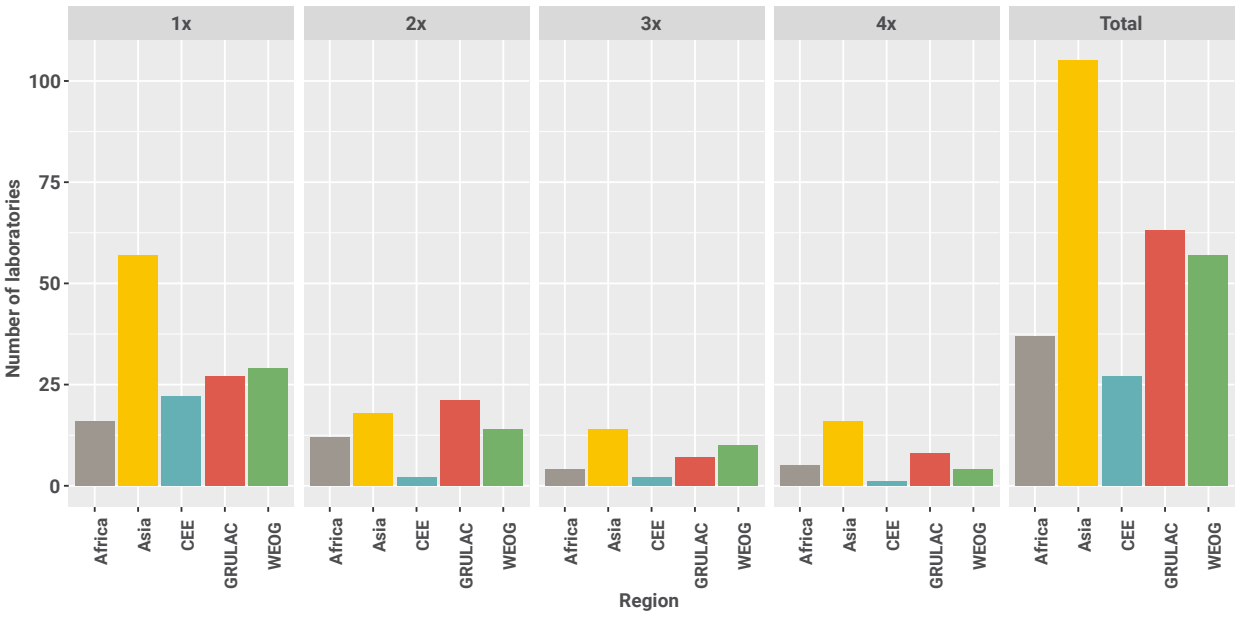


Figure 23: Number of laboratories by times of participation shown and colored by region

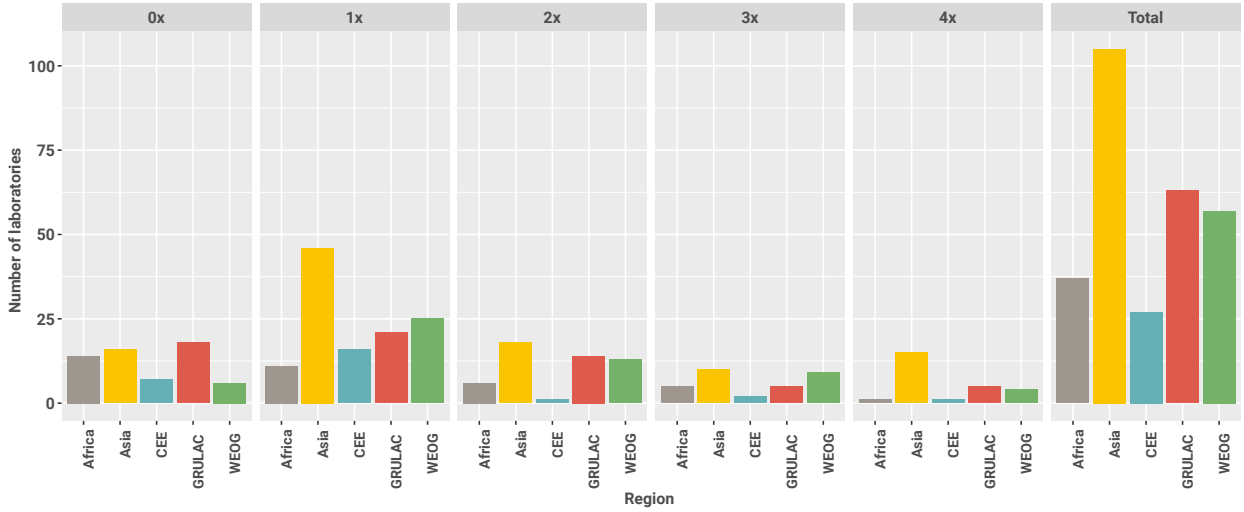


Figure 24: Number of results sets by laboratory grouped by times of delivery shown and colored by region

The reliability of the laboratories is shown by the difference between registration and results; however, this assessment cannot be done by direct comparison of the information in the two tables above but must be made laboratory by laboratory. The most "unreliable" laboratories are shown in Table 22; they had at least two more registrations than deliveries of results.

There was one laboratory from Africa that registered 3-times but never delivered results. Also, among the laboratories with four registrations, there were two laboratories, from Africa and GRULAC, that failed twice to deliver results. Seven laboratories registered twice without delivering results (four African, one CEE, and two GRULAC).

Table 22: Multiple registration: Number of laboratories that registered at least twice without delivering results in the same round

Region	Lab	Number results	Number registrations
Africa	L058	0x	3x
GRULAC	L063	2x	4x
Africa	L069	2x	4x
GRULAC	L045	0x	2x
Africa	L052	0x	2x
Africa	L095	0x	2x
Africa	L180	0x	2x
GRULAC	L215	0x	2x
CEE	L239	0x	2x
Africa	L245	0x	2x

5.3 Performance according to number of registrations

The impact of experience as number of registrations and number of results delivered in the ILs on the quality of the z-scores is shown in Figure 25. Due to the large number of z-scores, laboratories with 4x participation always have the largest share for each z-score; notably also for unsatisfactory (U) and inconsistent (I) z-scores. It can also be seen that laboratories did not always deliver in the round they registered.

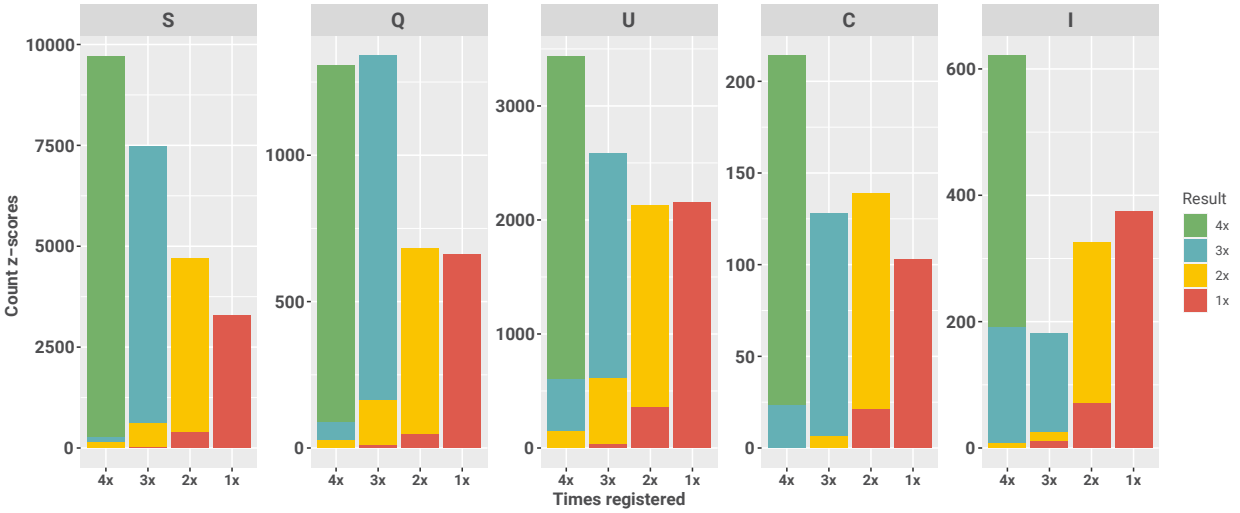


Figure 25: Summary by z-score by showing times of partcipation with number of registrations

5.4 Capacities and performance by laboratory

Figure 26 shows the number of POPs analytes that were tested in the interlaboratory assessments. Most POPs were tested in the test solutions of analytical standards (N=35), and 34 in the air extract (for toxaphene, there was no z-score assigned), 30 were for sediment and fish.

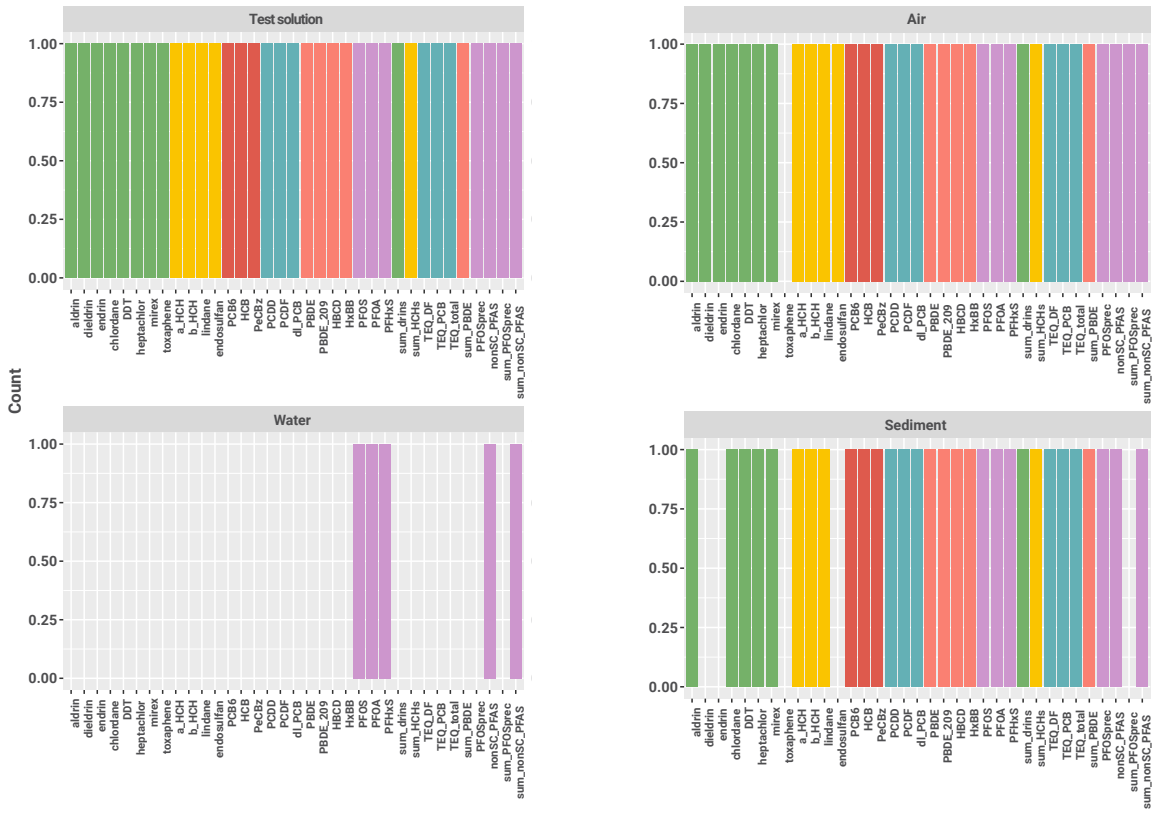


Figure 26: Overview of matrices and POPs to be tested and reported. The colors correspond the POP group: Green=OCPs, yellow=new OCPs, red= indPOPs, blue=dl-POPs, salmon=BFRs, violet=PFAS

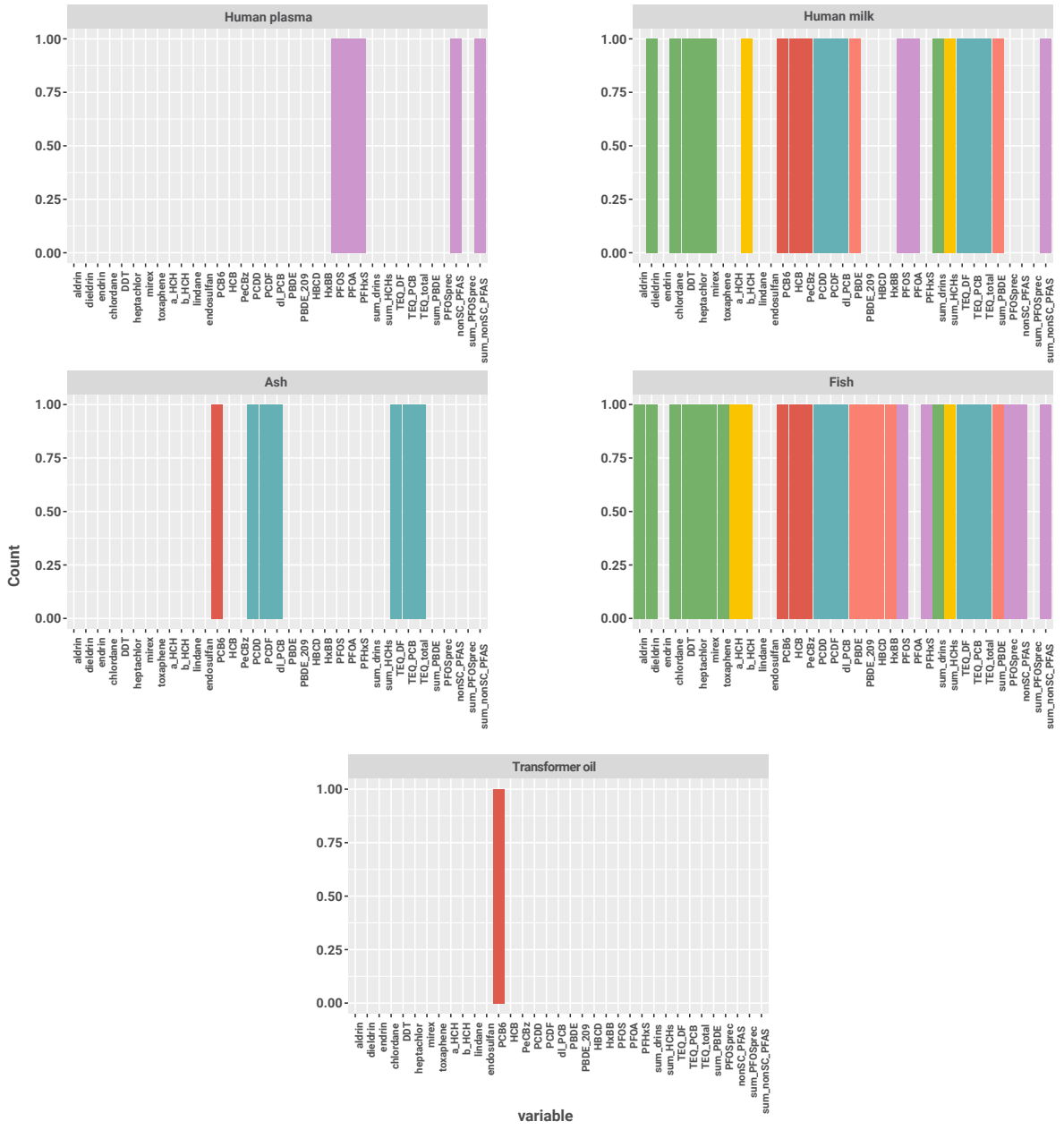


Figure 26: Overview of matrices and POPs to be tested and reported. The colors correspond the POP group: Green=OCPs, yellow=new OCPs, red= indPOPs, blue=dl-POPs, salmon=BFRs, violet=PFAS

In total, 204 combinations of POP and matrix have been assigned. There was one laboratory, L027, that obtained z-scores for 202 combinations. It did not analyze PCB6 in the transformer oil and not PFOS precursors in sediment, and so, was the laboratory with most experience. The other 22 laboratories that had more than half of the combinations analyzed were assessed together with L027. The results are shown in Figure 27.

Figure 27 shows that the majority of these laboratories analyzed a wide spectrum of POPs and matrices and have bars for all POPs on the x-axis; examples include especially L027 but also L011, L117. Other laboratories are specialized on biota samples, such as L030, L001, L124, L034 or L002. Laboratories, such as L072, L023, L128, L132 or L005 seem to be specialized on abiotic samples (and fish). Some laboratories, such as L126, L024, but also L101, did not/hardly analyze OCPs; laboratories L013 and L132 did not analyze PFAS, and laboratory L105 did not analyze dl-POPs.

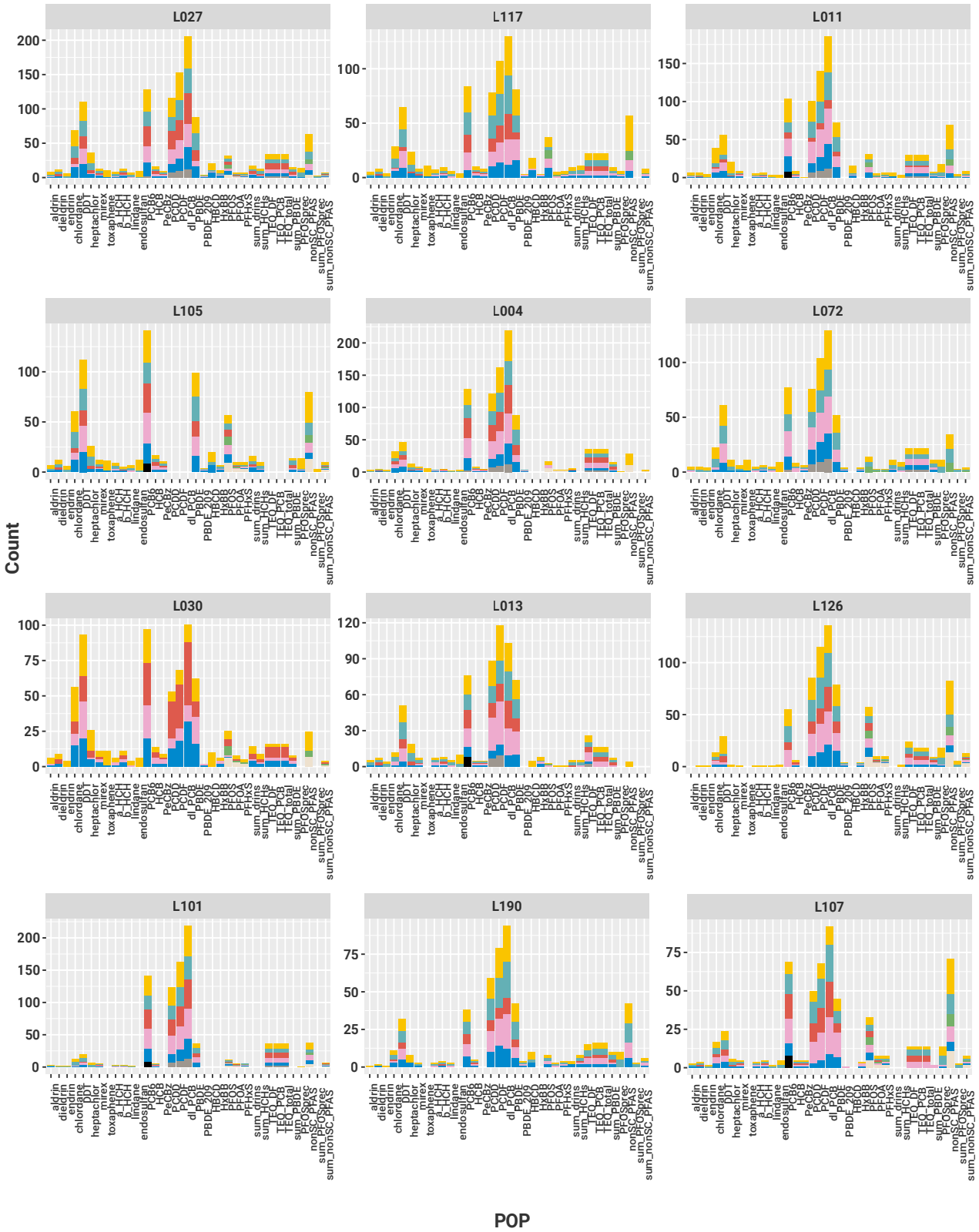


Figure 27: Overview of POPs and matrices analyzed by laboratory and z-scores obtained (top 23 laboratories)

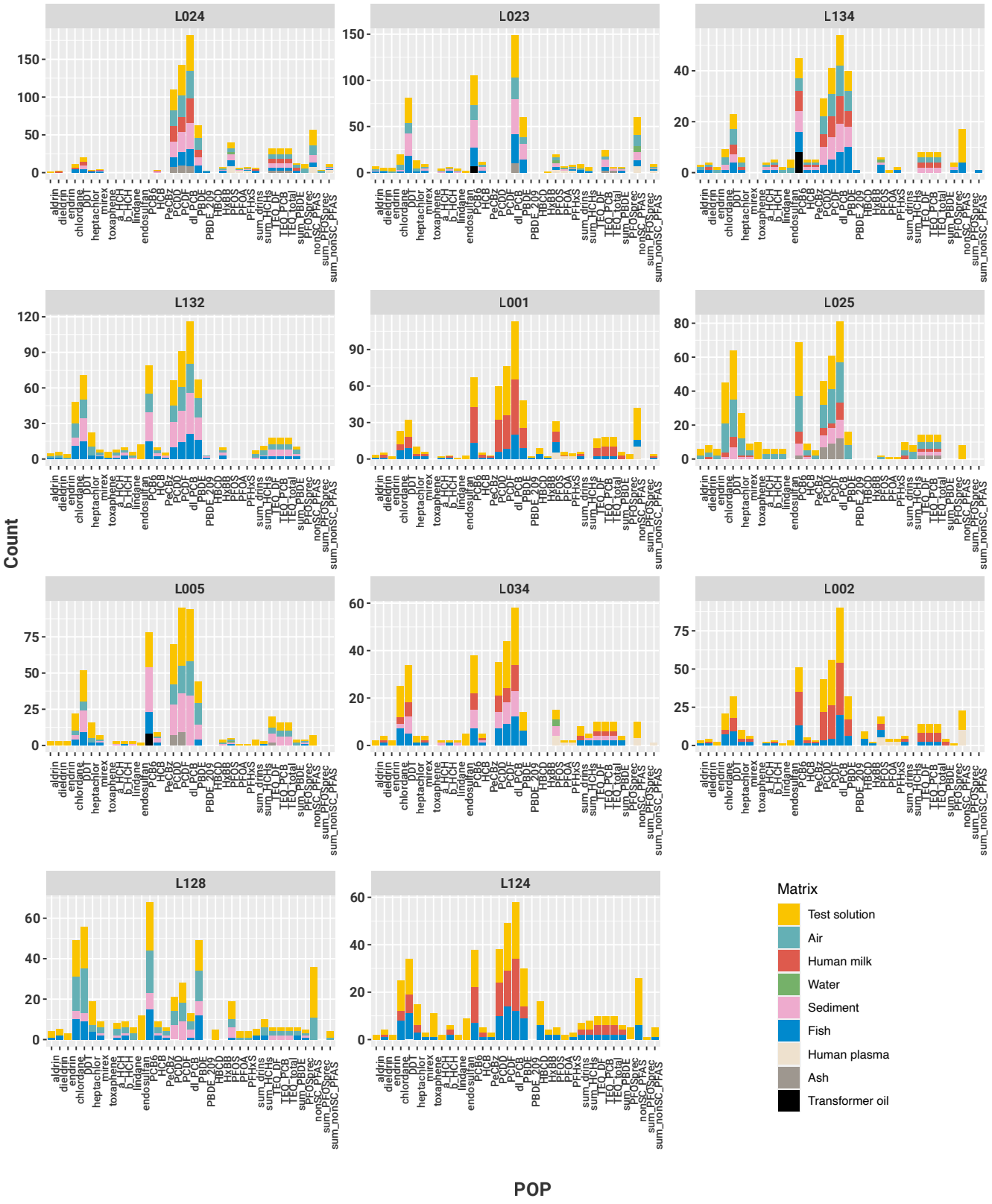


Figure 27: Overview of POPs and matrices analyzed by laboratory and z-scores obtained (top 23 laboratories)

5.5 Quality of the individual laboratories

An overview on the performance of the laboratories is visualized in the Appendix as Figure S 2. By comparing the scale of the green color with the scale of the red color in the stacked bars, some trends can be identified:

i. Some laboratories had good performances throughout several rounds of the interlaboratory assessment, *e.g.*, L001, L002, L004, L024, L026, L072, L125, L145;
ii. Some laboratories have increased the number of z-scores, indicating that capacity and experiences was build to include more POPs or mor matrices, *e.g.*, L022, L60, L072, L073, L105, L117, L126, L153;
iii. Unfortunately, some laboratories did not show improvement of performance; they either provided a large amount of unsatisfactory results across several rounds, *e.g.*, L018, L062, L106, L109, L163 or even increased the number of unsatisfactory results, *e.g.*, L007, L011 (at low level), L016, L018, L019, L023, L056;
iv. Some laboratories showed single unexpected results such as L003, which performed well in three rounds and almost completely failed in IL4 or L103 that had improved in IL3 and showed bad performance in IL4.

Finally, the laboratories that had obtained more than 200 z-scores in all the interlaboratory assessments were ranked according to their performance (i) according to the percentage of satisfactory results and (ii) according to the ratio of satisfactory to unsatisfactory results (S/U). 62 laboratories had achieved more than 200 z-scores and are considered further. There were three laboratories that had more than 1 000 z-scores (L027 – 1 335; L004 – 1 050, and L011 – 1 48). For details, see Table S 6.

For the overall assessment, there were 61% S results corresponding to a S/U ratio of 2.44. The ranking according to percentage of satisfactory results is shown in Table 23 and for the ratio S/U in Table 24. It can be seen that there are some laboratories with excellent performance (high percentages) but there were also laboratories that submitted many results and obtained many z-scores but having poor performance.

Table 23: Ranking of performance by laboratory according to percentage of satisfactory results within the laboratory's results

Lab	S	Lab	S	Lab	S
L025	91%	L072	74%	L007	56%
L125	90%	L153	73%	L035	55%
L034	89%	L011	72%	L115	55%
L027	88%	L126	71%	L195	53%
L242	87%	L101	71%	L013	50%
L094	86%	L107	70%	L065	50%
L037	85%	L128	69%	L060	49%
L124	83%	L003	67%	L073	49%
L001	82%	L104	64%	L135	48%
L012	82%	L148	64%	L016	43%
L029	81%	L190	64%	L053	40%
L002	81%	L112	64%	L041	38%
L008	80%	L134	64%	L023	37%
L017	80%	L147	64%	L050	33%
L024	79%	L156	63%	L102	31%
L030	78%	L031	62%	L019	25%
L145	78%	L173	62%	L091	17%
L005	78%	L022	60%	L163	6%
L117	78%	L105	59%	L062	5%
L004	76%	L015	58%		
L137	75%	L132	58%		

Laboratories with S/U ratio <1.0 had more unsatisfactory results than satisfactory results.

Table 24: Ranking of performance by laboratory according to ratio of satisfactory to unsatisfactory results

Lab	S/U	Lab	S/U	Lab	S/U
L125	42.67	L072	4.98	L003	2.36
L025	22.82	L137	4.71	L035	1.91
L027	21.29	L126	4.43	L013	1.81
L094	20.33	L128	4.24	L007	1.80
L034	17.56	L134	3.84	L060	1.68
L024	15.37	L011	3.82	L135	1.62
L037	12.93	L101	3.75	L065	1.51
L124	11.75	L156	3.63	L195	1.48
L107	10.14	L153	3.63	L073	1.41
L242	10.00	L147	3.49	L053	1.02
L017	9.43	L104	3.29	L023	0.89
L001	8.81	L105	3.18	L050	0.80
L145	8.23	L031	3.17	L016	0.78
L030	7.99	L132	2.83	L041	0.72
L002	7.80	L112	2.67	L102	0.68
L012	7.72	L148	2.59	L019	0.42
L117	7.18	L015	2.59	L091	0.32
L029	6.96	L173	2.58	L062	0.10
L005	6.87	L115	2.56	L163	0.07
L008	6.08	L190	2.55		
L004	5.82	L022	2.55		

6

Conclusions

Over about ten years and four rounds, the UNEP interlaboratory assessments have gained international reputation. A total of 289 laboratories from 82 countries participated, with up to eight test solutions and eight test matrices. The interlaboratory assessments were larger than other proficiency test in terms of chemicals to be analyzed and matrices offered. Conclusions and lessons learned can be drawn on the following:

6.1 Participation with view on registration and delivery of results

289 laboratories from 82 countries responded once or more frequently to the invitation to participate in one of the rounds of the interlaboratory assessment. Results, as datasets of amounts identified and quantified in the test samples, were received from 228 laboratories in 72 countries. Thus, there were 61 laboratories and 10 countries that did not fulfil the expectations for their commitments.

There were 532 registrations by laboratories but only 76% delivered (420); thus, overall almost one quarter of the laboratories failed. We consider this a strong indicator that there were many laboratories too ambitious; they overestimated their analytical capacity. Assessing the IL1 had 103 laboratories registered and 82 sets of results (80% delivered), for IL2, the ratio was 105 to 89 or 78%, and IL4 had 148 registrations toward 116 deliveries, corresponding to 78%. Almost no improvement was made from IL1 to IL4, so that in general about 20% (or more) of the efforts failed.

The most stable participation is from the Asian region with 15 laboratories submitting results in four rounds and another nine with 3-times delivery of results. The laboratories from the WEOG region had lower participation but is increasing: 4 laboratories provided results 4-times and 9 laboratories 3-times. This finding may be due to two criteria: (i) the first round was targeted towards developing country regions laboratories and did not include so many WEOG laboratories; (ii) according to GEF rules, the laboratories from WEOG had to pay for their participation whereas for the developing country regions, the participation fee was covered from external funds (mostly GEF projects).

Reasons that laboratories were unable to deliver results although they had registered for participation included:

- Most registered laboratories were unable to participate as they did not manage to analyse the samples with their equipment or method. Very often, the analytical equipment, especially the detectors, were not operational.
- Many laboratories informed that they did not have analytical standards for identification and quantification.
- In a few cases, test samples could not be shipped from the coordinating laboratory to the recipient laboratory or the recipient laboratory was not able to accept the samples. Reasons were wrong address details, recipient laboratory did not respond to announcement of arrival of the package, lack of permit for import of the materials (customs' issues).
- In very few cases, the sample was destroyed during transport; *e.g.*, leakage or break of the transport vessel.
- There were some limitations with the analyses of certain POPs regionally, for example the analyses of PFAS in Africa and GRULAC, additional to the

analyses of brominated flame retardants in Africa.

- There is a need to continue working on improving the quality of POPs analysis worldwide. This is evident considering that during each meeting of the COPs new POPs are included to the Convention, and there are already listed POPs, such as polychlorinated naphthalenes and chain chlorinated paraffins, which were not included in the interlaboratory assessment.

- Regular interlaboratory assessment including different combinations of POPs and matrices are for control of assessment. The routine analyses of the laboratories should be the same used for generating and reporting results.

- To improve and ensure better POPs analyses, a good quality of laboratories is needed including the instrumentation as well as all the aspects of extraction, clean-up steps, materials, consumables (certified standards, high-purity solvent, and high-quality gases) and skilled personnel. It is important that efforts are made to strive for gender parity within the laboratories. It is also important throughout normal operations to apply self-control, quality assurance measures and quality controls charts.

- To ensure sustainability and maintenance of the infrastructure and instrument the laboratories need to follow a business plan of routine POPs analyses.

- Information generated through high quality analysis of POPs represents a valuable resource for both policy makers and researchers worldwide. It creates an opportunity to further explore complex issues such as gender and age-differentiated windows of exposure and the relationship between POPs and vulnerable groups.

©UNEP/3rd Interlab Assessment Result Workshop 2017, Beijing, China

7

References

Abalos, M., Abad, E., van Leeuwen, S.P.J., Lindstrom, G., Fiedler, H. *et al.* (2013). Results for PCDD/PCDF and dl-PCBs in the First Round of UNEPs Biennial Global Inter laboratory Assessment on Persistent Organic Pollutants. *TrAC-Trends in Analytical Chemistry* 46, 98-109. https://doi.org/10.1016/j.trac.2012.11.003.

Cofino, W.P., van Stokkum, I.H.M., van Steenwijk, J. and Wells, D.E. (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. https://doi.org/10.1016/j.aca.2004.11.008.

de Boer, J., Leslie, H., van Leeuwen, S.P., Wegener, J.W., van Bavel, B., Lindstrom, G. *et al.* (2008). United Nations Environment Programme Capacity Building Pilot Project--training and interlaboratory study on persistent organic pollutant analysis under the Stockholm Convention. *Analytica Chimica Acta* 617(1-2), 208-215. https://dx.doi.org/10.1016/j.aca.2008.01.081.

de Boer, J., van der Veen, I. and Fiedler, H. (2022). Global interlaboratory assessments on PCBs, organochlorine pesticides and brominated flame retardants in various environmental matrices 2017/2019. *Chemosphere* 295, 133991. https://doi.org/10.1016/j.chemosphere.2022.133991.

Fiedler, H., de Boer, J. and van Bavel, B. (2016). Assessment of results for the 2nd interlaboratory study of POPs laboratories. *Organohalogen Compounds* 78, 777–780. https://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-66223.

Fiedler, H., van der Veen, I. and de Boer, J. (2020). Global interlaboratory assessments of perfluoroalkyl substances under the Stockholm Convention on persistent organic pollutants. *Trac-Trends in Analytical Chemistry* 124, 115459. https://doi.org/10.1016/j.trac.2019.03.023.

Fiedler, H., van der Veen, I. and de Boer, J. (2022a). Interlaboratory assessments for dioxin-like POPs (2016/2017 and 2018/2019). *Chemosphere* 288, 132449. https://doi.org/10.1016/j.chemosphere.2021.132449.

Fiedler, H., van der Veen, I. and de Boer, J. (2022b). Assessment of four rounds of interlaboratory tests within the UNEP-coordinated POPs projects. *Chemosphere* 288, 132441. https://doi.org/10.1016/j.chemosphere.2021.132441.

Molenaar, J., Cofino, W.P. and Torfs, P.J.J.F. (2018). Efficient and robust analysis of interlaboratory studies. *Chemometrics and Intelligent Laboratory Systems* 175, 65-73. https://doi.org/10.1016/j.chemolab.2018.01.003.

Thompson, M., Ellison, S. and Wood, R. (2006). The international harmonized protocol for the proficiency testing of analytical chemistry laboratories: (IUPAC technical report). *Pure and Applied Chemistry* 78, 145-196. https://doi.org/10.1351/pac200678010145.

United Nations Environment Programme (2005). *International Intercalibration Studies: A Global QA/QC Tool for the Analysis of POPs under the Stockholm Convention.* Geneva.

United Nations Environment Programme (2012). *Bi-ennial Global Interlaboratory Assessment on Persistent Organic Pollutants – First Round 2010/2011.* https://wedocs.unep.org/20.500.11822/31374.

United Nations Environment Programme (2015). *Bi-ennial Global Interlaboratory Assessment on Persistent Organic Pollutants - 2nd Round 2012-2013.* https://wedocs.unep.org/bitstream/handle/20.500.11822/31377/POPs_IA_2nd_Round.pdf?sequence=1&isAllowed=y.

United Nations Environment Programme (2017). *Bi-ennial Global Interlaboratory Assessment on Persistent Organic Pollutants – Third Round 2016/2017.* Geneva. https://wedocs.unep.org/handle/20.500.11822/21743.

United Nations Environment Programme (2019). *Global Chemicals Outlook II: From Legacies to Innovative Solutions - Implementing the 2030 Agenda for Sustainable Development.* Geneva. https://wedocs.unep.org/20.500.11822/28113.

United Nations Environment Programme (2021a). *Guidance on the Global Monitoring Plan for Persistent Organic Pollutants.* 21 April, UNEP/POPS/COP.10/INF/42. Geneva: United Nations Environment Programme, Secretariat of the Basel, Rotterdam and Stockholm Conventions. http://chm.pops.int/TheConvention/ConferenceoftheParties/Meetings/COP10/tabid/8397/Default.aspx.

United Nations Environment Programme (2021b). *Bi-ennial Global Interlaboratory Assessment on Persistent Organic Pollutants – Fourth Round 2018/2019.* Geneva. https://wedocs.unep.org/bitstream/handle/20.500.11822/30923/R4-Interlab-POPs.pdf?sequence=1&isAllowed=y.

van Bavel, B., van der Veen, I., Nilsson, H., de Boer, J., and Fiedler, H. (2014). Results from UNEPs 2nd global interlaboratory assessment: dioxinlike POPs. *Organohalogen Compounds* 76, 1632-1635. http://oru.diva-portal.org/smash/record.jsf?pid=diva2%3A1194096&dswid=-1208.

van der Veen, I., de Boer, J. and Fiedler, H. (2017). Bi-ennial global interlaboratory assessment on persistent organic pollutants – Third Round 2016/2017, organochlorine pesticides, PCBs and brominated flame retardants. *Organohalogen Compounds* 79, 575-578. https://www.diva-portal.org/smash/get/diva2:1284933/FULLTEXT01.pdf.

van der Veen, I., Fiedler, H. and de Boer, J. (2023). Assessment of the per- and polyfluoroalkyl substances analysis under the Stockholm Convention - 2018/2019. *Chemosphere* 313, 137549. https://doi.org/10.1016/j.chemosphere.2022.137549.

van Leeuwen, S.P.J., de Boer, J., van Leeuwen, S.P.J., and van Bavel, B. (2013). First worldwide UNEP interlaboratory study on persistent organic pollutants (POPs), with data on polychlorinated biphenyls and organochlorine pesticides. *TrAC Trends in Analytical Chemistry* 46, 110-117. https://doi.org/10.1016/j.trac.2012.12.020.

8

Appendix

Table S 1: Nomenclature and grouping of determinands/POP species in the interlaboratory assessments by POP group

POP_group	POP_subgroup	POP	POP_species
OPCs	drins	aldrin	aldrin
		dieldrin	dieldrin
		endrin	endrin
		endrin	endrin_ketone[1]
		sum_drins	drins_LB
		sum_drins	drins_UB
	chlordane	chlordane	*cis*-chlordane
		chlordane	*trans*-chlordane
		chlordane	oxychlordane
		chlordane	cis_nonachlor
		chlordane	trans_nonachlor
		chlordane	chlordane_LB
		chlordane	chlordane_UB
	DDT	DDT	op_DDT
		DDT	op_DDD
		DDT	op_DDE
		DDT	pp_DDT
		DDT	pp_DDD
		DDT	pp_DDE
		DDT	DDT_LB
		DDT	DDT_UB
	chlordecone	chlordecone	chlordecone[2]
	endosulfan	endosulfan	a_endosulfan
		endosulfan	b_endosulfan
		endosulfan	endosulfan sulfate
		endosulfan	endosulfan_LB
		endosulfan	endosulfan_UB

[1] Without z-score throughout all rounds of the interlaboratory assessment
[2] Without z-score throughout all rounds of the interlaboratory assessment

POP_group	POP_subgroup	POP	POP_species
OPCs	heptachlor	heptachlor	heptachlor
		heptachlor	cis_hepo
		heptachlor	trans_hepo
		heptachlor	heptachlor_LB
		heptachlor	heptachlor_UB
	HCHs	α-HCH	α_HCH
		β-HCH	β_HCH
		lindane	lindane
		sum_HCHs	HCHs_LB
		sum_HCHs	HCHs_UB
	mirex	mirex	mirex
	toxaphene	toxaphene	Parlar_26
		toxaphene	Parlar_50
		toxaphene	Parlar_62
		toxaphene	toxaphene_LB
		toxaphene	toxaphene_UB
indPOP	HCB	HCB	HCB
	PCB6	PCB6	PCB_28
		PCB6	PCB_52
		PCB6	PCB_101
		PCB6	PCB_138
		PCB6	PCB_153
		PCB6	PCB_180
		PCB6	PCB(6)_LB
		PCB6	PCB(6)_UB
	PeCBz	PeCBz	PeCBz
	HCBD	HCBD	HCBD
dl_POPs	PCDD	PCDD	Cl4DD
		PCDD	Cl5DD
		PCDD	Cl6DD1
		PCDD	Cl6DD2
		PCDD	Cl6DD3
		PCDD	Cl7DD
		PCDD	OCDD

POP_group	POP_subgroup	POP	POP_species
dl_POPs	PCDF	PCDF	Cl4DF
		PCDF	Cl5DF1
		PCDF	Cl5DF2
		PCDF	Cl6DF1
		PCDF	Cl6DF2
		PCDF	Cl6DF3
		PCDF	Cl6DF4
		PCDF	Cl7DF1
		PCDF	Cl7DF2
		PCDF	OCDF
	PCDD/PCDF	TEQ_DF	TEQ(DF)_LB
		TEQ_DF	TEQ(DF)_UB
	dl_PCB	dl_PCB	PCB_77
		dl_PCB	PCB_81
		dl_PCB	PCB_105
		dl_PCB	PCB_114
		dl_PCB	PCB_118
		dl_PCB	PCB_123
		dl_PCB	PCB_126
		dl_PCB	PCB_156
		dl_PCB	PCB_157
		dl_PCB	PCB_167
		dl_PCB	PCB_169
		dl_PCB	PCB_189
		TEQ_PCB	TEQ(PCB)_LB
		TEQ_PCB	TEQ(PCB)_UB
	dl_POPs	TEQ_total	TEQ(total)_LB
		TEQ_total	TEQ(total)_UB
BFR	PBDE	PBDE	PBDE_17
		PBDE	PBDE_28
		PBDE	PBDE_47
		PBDE	PBDE_99
		PBDE	PBDE_100
		PBDE	PBDE_153
		PBDE	PBDE_154
		PBDE	PBDE_183

POP_group	POP_subgroup	POP	POP_species
		PBDE_209	PBDE_209
		sum_PBDE	PBDE_LB
		sum_PBDE	PBDE_UB
	HxBB	HxBB	PBB_153
	HBCD	HBCD	a_HBCD
		HBCD	b_HBCD
		HBCD	g_HBCD
		HBCD	HBCD_LB
		HBCD	HBCD_UB
PFAS	PFOS	PFOS	L_PFOS
		PFOS	br_PFOS
		PFOS	PFOS(tot)_LB
		PFOS	PFOS(tot)_UB
	PFOA	PFOA	PFOA
	PFHxS	PFHxS	L_PFHxS
	PFOSprec	PFOSprec	FOSA
		PFOSprec	EtFOSA
		PFOSprec	EtFOSE
		PFOSprec	MeFOSA
		PFOSprec	MeFOSE
		sum_PFOSprec	PFOSprec_LB
		sum_PFOSprec	PFOSprec_UB
	nonSC_PFAS	nonSC_PFAS	PFBS, L_PFBS
		nonSC_PFAS	L_PFHpS
		nonSC_PFAS	L_PFDS
		nonSC_PFAS	PFBA
		nonSC_PFAS	PFPeA
		nonSC_PFAS	PFHxA
		nonSC_PFAS	PFHpA
		nonSC_PFAS	PFNA
		nonSC_PFAS	PFDA
		nonSC_PFAS	PFUnDA
		nonSC_PFAS	PFDoDA
		nonSC_PFAS	PFTrDA
		nonSC_PFAS	PFTeDA
		nonSC_PFAS	FTSA_62
		sum_nonSC_PFAS	PFCA+PFSA_LB
		sum_nonSC_PFAS	PFCA+PFSA_UB

Table S 2: Overview on registration and z-scores assigned *(i.e.,* delivering results) by laboratory and round

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L001	Y	Y	Y	Y	Y	Y	Y	Y
L002	Y	Y	Y		Y	Y	Y	notR
L003	Y	Y	Y	Y	Y	Y	Y	Y
L004	Y	Y	Y	Y	Y	Y	Y	Y
L005	Y	Y	Y	Y	Y	Y	Y	Y
L006	Y	Y			Y	Y	notR	notR
L007	Y	Y	Y		Y	Y	Y	notR
L008	Y	Y	Y	Y	Y	Y	Y	Y
L009	Y	Y	Y		Y	Y	Y	notR
L010	Y	Y			Y	Y	notR	notR
L011	Y	Y	Y	Y	Y	Y	Y	Y
L012	Y	Y	Y		Y	Y	Y	notR
L013	Y	Y	Y	Y	Y	Y	Y	Y
L014	Y	Y	Y		Y	Y	N	notR
L015	Y	Y	Y		Y	Y	Y	notR
L016	Y	Y	Y	Y	Y	Y	Y	Y
L017	Y	Y	Y	Y	Y	Y	Y	Y
L018	Y	Y	Y	Y	Y	Y	Y	Y
L019	Y	Y		Y	Y	N	notR	Y
L020	Y	Y	Y		Y	Y	Y	notR
L021	Y				N	notR	notR	notR
L022	Y	Y	Y	Y	Y	Y	Y	Y
L023	Y	Y	Y	Y	Y	Y	Y	Y
L024	Y	Y	Y	Y	Y	Y	Y	Y
L025	Y	Y	Y	Y	Y	Y	Y	Y
L026	Y				Y	notR	notR	notR
L027	Y	Y	Y	Y	Y	Y	Y	Y
L028	Y				Y	notR	notR	notR
L029	Y	Y			Y	Y	notR	notR
L030	Y	Y	Y	Y	Y	Y	Y	Y
L031	Y	Y		Y	Y	Y	notR	Y
L032	Y	Y			Y	Y	notR	notR
L033	Y				Y	notR	notR	notR
L034	Y		Y		Y	notR	Y	notR
L035	Y	Y	Y	Y	Y	Y	Y	Y
L036	Y				Y	notR	notR	notR
L037	Y		Y	Y	Y	notR	Y	Y
L038	Y				Y	notR	notR	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L040	Y				N	notR	notR	notR
L041	Y	Y	Y		Y	Y	N	notR
L042	Y	Y	Y		Y	Y	N	notR
L043	Y		Y	Y	Y	notR	Y	Y
L044	Y				N	notR	notR	notR
L045	Y		Y		N	notR	N	notR
L046	Y				Y	notR	notR	notR
L047	Y	Y			Y	N	notR	notR
L048	Y	Y			Y	N	notR	notR
L049	Y		Y	Y	Y	notR	Y	N
L050	Y	Y	Y	Y	Y	Y	Y	Y
L051	Y				Y	notR	notR	notR
L052	Y			Y	N	notR	notR	N
L053	Y		Y	Y	Y	notR	Y	Y
L054	Y				Y	notR	notR	notR
L055	Y				N	notR	notR	notR
L056	Y	Y	Y	Y	Y	N	Y	Y
L057	Y				Y	notR	notR	notR
L058	Y	Y		Y	N	N	notR	N
L059	Y				N	notR	notR	notR
L060	Y	Y	Y	Y	Y	Y	Y	Y
L061	Y	Y	Y	Y	Y	Y	Y	Y
L062	Y	Y	Y	Y	Y	Y	N	Y
L063	Y	Y	Y	Y	Y	Y	N	N
L064	Y	Y	Y		Y	Y	Y	notR
L065	Y	Y	Y	Y	Y	Y	Y	Y
L066	Y				N	notR	notR	notR
L067	Y	Y		Y	Y	N	notR	N
L068	Y	Y	Y	Y	Y	Y	N	Y
L069	Y	Y	Y	Y	Y	N	Y	N
L070	Y	Y			N	Y	notR	notR
L071	Y			Y	N	notR	notR	Y
L072	Y		Y	Y	Y	notR	Y	Y
L073	Y	Y	Y		Y	Y	Y	notR
L074	Y	Y	Y	Y	Y	Y	Y	N
L075	Y				Y	N	notR	notR
L076	Y	Y	Y		Y	N	Y	notR
L077	Y				Y	notR	notR	notR
L079	Y				Y	notR	notR	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L080	Y	Y	Y	Y	Y	N	Y	Y
L081	Y	Y	Y		Y	Y	N	notR
L082	Y			Y	N	notR	notR	Y
L083	Y		Y	Y	N	notR	N	Y
L084	Y				Y	notR	notR	notR
L085	Y				Y	notR	notR	notR
L086	Y	Y	Y	Y	Y	N	Y	Y
L087	Y	Y		Y	Y	Y	notR	N
L088	Y				N	notR	notR	notR
L089	Y				N	notR	notR	notR
L090	Y				Y	notR	notR	notR
L091	Y	Y	Y	Y	Y	Y	Y	Y
L092	Y		Y		N	notR	Y	notR
L093	Y		Y		N	notR	Y	notR
L094	Y	Y	Y	Y	Y	Y	Y	Y
L095	Y	Y			N	N	notR	notR
L096	Y			Y	Y	notR	notR	N
L097	Y	Y			N	Y	notR	notR
L098	Y				Y	notR	notR	notR
L099	Y				N	notR	notR	notR
L100	Y				Y	notR	notR	notR
L101	Y	Y	Y	Y	Y	Y	Y	Y
L102	Y		Y	Y	Y	notR	Y	Y
L103	Y	Y	Y	Y	Y	Y	Y	Y
L104	Y		Y	Y	Y	notR	Y	Y
L105	Y	Y	Y	Y	Y	Y	Y	Y
L106		Y	Y	Y	notR	Y	Y	Y
L107		Y		Y	notR	Y	notR	Y
L109		Y			notR	N	notR	notR
L110		Y			notR	Y	notR	notR
L111		Y			notR	Y	notR	notR
L112		Y			notR	Y	notR	notR
L113			Y		notR	notR	Y	notR
L114		Y			notR	Y	notR	notR
L115		Y	Y	Y	notR	Y	Y	Y
L116		Y			notR	Y	notR	notR
L117		Y	Y	Y	notR	Y	Y	Y
L119		Y			notR	Y	notR	notR
L120		Y	Y		notR	Y	Y	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L121		Y	Y		notR	Y	Y	notR
L122		Y			notR	Y	notR	notR
L123			Y	Y	notR	notR	Y	Y
L124			Y	Y	notR	notR	Y	Y
L125		Y	Y	Y	notR	Y	Y	Y
L126		Y	Y	Y	notR	Y	Y	Y
L127		Y	Y		notR	N	Y	notR
L128		Y	Y	Y	notR	Y	Y	Y
L129		Y	Y		notR	Y	Y	notR
L130		Y		Y	notR	Y	notR	Y
L131		Y			notR	Y	notR	notR
L132		Y	Y	Y	notR	Y	Y	Y
L133		Y			notR	N	notR	notR
L134		Y		Y	notR	Y	notR	Y
L135		Y			notR	Y	notR	notR
L136		Y		Y	notR	Y	notR	N
L137		Y	Y		notR	Y	Y	notR
L139		Y	Y		notR	Y	Y	notR
L140		Y			notR	Y	notR	notR
L141		Y			notR	Y	notR	notR
L142		Y			notR	Y	notR	notR
L143		Y			notR	Y	notR	notR
L144		Y			notR	N	notR	notR
L145		Y	Y	Y	notR	Y	Y	Y
L146		Y			notR	Y	notR	notR
L147		Y	Y	Y	notR	Y	Y	N
L148		Y		Y	notR	Y	notR	Y
L149		Y	Y	Y	notR	Y	Y	Y
L150		Y			notR	N	notR	notR
L151		Y			notR	Y	notR	notR
L152		Y	Y		notR	Y	Y	notR
L153		Y	Y	Y	notR	Y	Y	Y
L154		Y			notR	Y	notR	notR
L155		Y	Y		notR	Y	Y	notR
L156			Y	Y	notR	notR	Y	Y
L157			Y		notR	notR	Y	notR
L158			Y		notR	notR	Y	notR
L159			Y		notR	notR	N	notR
L160			Y		notR	notR	N	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L161			Y	Y	notR	notR	Y	Y
L162			Y		notR	notR	N	notR
L163			Y	Y	notR	notR	Y	Y
L164			Y	Y	notR	notR	Y	Y
L165			Y		notR	notR	N	notR
L166			Y	Y	notR	notR	Y	Y
L167			Y	Y	notR	notR	Y	Y
L168			Y		notR	notR	N	notR
L169			Y		notR	notR	Y	notR
L170			Y		notR	notR	Y	notR
L171			Y		notR	notR	Y	notR
L172			Y		notR	notR	N	notR
L173			Y	Y	notR	notR	Y	Y
L174			Y		notR	notR	N	notR
L175			Y		notR	notR	Y	notR
L176			Y	Y	notR	notR	Y	N
L177			Y		notR	notR	N	notR
L178			Y		notR	notR	N	notR
L179			Y	Y	notR	notR	Y	N
L180			Y	Y	notR	notR	N	N
L182			Y	Y	notR	notR	Y	Y
L183			Y	Y	notR	notR	Y	N
L184			Y		notR	notR	N	notR
L185			Y		notR	notR	Y	notR
L186			Y	Y	notR	notR	Y	Y
L187			Y	Y	notR	notR	Y	N
L188			Y	Y	notR	notR	Y	Y
L189			Y	Y	notR	notR	Y	Y
L190			Y	Y	notR	notR	Y	Y
L191			Y	Y	notR	notR	Y	Y
L192			Y		notR	notR	N	notR
L194			Y	Y	notR	notR	Y	Y
L195			Y	Y	notR	notR	Y	Y
L196			Y	Y	notR	notR	Y	Y
L197			Y		notR	notR	Y	notR
L198			Y		notR	notR	Y	notR
L199			Y		notR	notR	N	notR
L200			Y		notR	notR	N	notR
L206			Y		notR	notR	Y	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L207			Y	Y	notR	notR	Y	notR
L208			Y	Y	notR	notR	Y	Y
L209			Y		notR	notR	Y	notR
L210			Y		notR	notR	N	notR
L211			Y		notR	notR	Y	notR
L212			Y		notR	notR	N	notR
L213			Y		notR	notR	Y	notR
L214			Y		notR	notR	N	notR
L215			Y	Y	notR	notR	N	N
L216			Y		notR	notR	N	notR
L219			Y		notR	notR	Y	notR
L220			Y		notR	notR	Y	notR
L221			Y		notR	notR	Y	notR
L222			Y		notR	notR	N	notR
L223			Y		notR	notR	Y	notR
L224			Y	Y	notR	notR	Y	Y
L225			Y		notR	notR	N	notR
L226			Y	Y	notR	notR	Y	Y
L227			Y		notR	notR	Y	notR
L228			Y		notR	notR	N	notR
L229			Y	Y	notR	notR	Y	Y
L230			Y		notR	notR	Y	notR
L231			Y		notR	notR	Y	notR
L232			Y		notR	notR	Y	notR
L233			Y	Y	notR	notR	Y	Y
L234			Y		notR	notR	Y	notR
L235			Y		notR	notR	Y	notR
L236			Y		notR	notR	N	notR
L237			Y		notR	notR	Y	notR
L238			Y	Y	notR	notR	Y	Y
L239			Y	Y	notR	notR	N	N
L240			Y		notR	notR	Y	notR
L241			Y		notR	notR	N	notR
L242			Y	Y	notR	notR	Y	Y
L243			Y		notR	notR	N	notR
L244			Y	Y	notR	notR	N	Y
L245			Y	Y	notR	notR	N	N
L246			Y		notR	notR	N	notR
L247			Y		notR	notR	N	notR

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L248			Y		notR	notR	N	notR
L249			Y		notR	notR	N	notR
L250			Y		notR	notR	Y	notR
L251			Y		notR	notR	N	notR
L252			Y		notR	notR	Y	notR
L253			Y		notR	notR	Y	notR
L254			Y		notR	notR	Y	notR
L255			Y	Y	notR	notR	Y	Y
L256			Y		notR	notR	Y	notR
L257			Y		notR	notR	Y	notR
L258			Y		notR	notR	Y	notR
L259				Y	notR	notR	notR	Y
L260				Y	notR	notR	notR	N
L261				Y	notR	notR	notR	Y
L262				Y	notR	notR	notR	Y
L263				Y	notR	notR	notR	N
L264				Y	notR	notR	notR	Y
L265			Y	Y	notR	notR	N	Y
L266				Y	notR	notR	notR	Y
L267				Y	notR	notR	notR	Y
L268				Y	notR	notR	notR	Y
L269				Y	notR	notR	notR	Y
L270				Y	notR	notR	notR	Y
L271				Y	notR	notR	notR	Y
L272				Y	notR	notR	notR	Y
L273				Y	notR	notR	notR	N
L274				Y	notR	notR	notR	N
L275				Y	notR	notR	notR	Y
L276				Y	notR	notR	notR	Y
L278				Y	notR	notR	notR	Y
L279				Y	notR	notR	notR	Y
L281				Y	notR	notR	notR	N
L282				Y	notR	notR	notR	Y
L283				Y	notR	notR	notR	N
L284				Y	notR	notR	notR	Y
L286				Y	notR	notR	notR	Y
L287				Y	notR	notR	notR	Y
L288				Y	notR	notR	notR	Y
L289				Y	notR	notR	notR	Y

Lab	Reg_IL1	Reg_IL2	Reg_IL3	Reg_IL4	Res_IL1	Res_IL2	Res_IL3	Res_IL4
L290				Y	notR	notR	notR	Y
L291				Y	notR	notR	notR	Y
L292				Y	notR	notR	notR	N
L293				Y	notR	notR	notR	Y
L294				Y	notR	notR	notR	N
L295				Y	notR	notR	notR	N
L296				Y	notR	notR	notR	Y
L297				Y	notR	notR	notR	Y
L298				Y	notR	notR	notR	Y
L299				Y	notR	notR	notR	N
L300				Y	notR	notR	notR	N
L301				Y	notR	notR	notR	Y
L302				Y	notR	notR	notR	Y
L303				Y	notR	notR	notR	Y
L304				Y	notR	notR	notR	Y
L305				Y	notR	notR	notR	N
L306				Y	notR	notR	notR	Y

Legend:

Y	Laboratories that registered and obtained at least one z-score in any of the four interlaboratory assessments.
notR	Laboratories that did not register.
N	Laboratories that registered but either not delivered results or did not obtain any z-score.
	Laboratory did not participate in the respective round.

Table S 3: Summary of laboratories *per* region and participation in round of the IL. Reg=Registration; Result=Laboratory delivered result

Round	IL1		IL2		IL3		IL4		Sum	
Particip	Reg	Result	Reg	Result	Reg	Result	Reg	Result	Reg	Result
Africa	17	10	12	5	19	14	24	13	72	42
	L052, L053, L056, L058, L066, L067, L069, L074, L077, L079, L082, L086, L089, L091, L095, L097, L100, L106, L127, L155, L163, L171, L177, L180, L186, L191, L196, L199, L236, L245, L270, L273, L274, L281, L282, L295, L303									
Asia	38	33	45	42	68	53	48	44	199	172
	L001, L002, L003, L004, L005, L006, L007, L008, L009, L010, L011, L012, L013, L014, L015, L016, L017, L018, L019, L020, L021, L022, L023, L025, L027, L028, L030, L032, L038, L040, L041, L042, L059, L064, L068, L073, L076, L093, L111, L114, L119, L120, L121, L122, L123, L137, L140, L144, L148, L151, L153, L154, L156, L157, L158, L159, L166, L167, L169, L173, L178, L185, L187, L190, L198, L200, L206, L207, L221, L222, L225, L226, L227, L234, L235, L243, L244, L247, L248, L249, L250, L251, L252, L253, L254, L258, L259, L261, L266, L268, L269, L271, L272, L278, L279, L284, L293, L296, L297, L299, L300, L301, L302, L304, L306									
CEE	3	3	4	4	23	16	6	5	36	28
	L037, L046, L050, L112, L113, L116, L149, L162, L165, L168, L170, L172, L174, L175, L197, L209, L219, L220, L230, L231, L232, L233, L237, L239, L240, L241, L289									
GRULAC	32	23	14	11	39	25	37	25	122	84
	L043, L044, L045, L047, L048, L049, L051, L054, L055, L057, L060, L061, L062, L063, L065, L070, L071, L072, L080, L081, L083, L084, L085, L087, L088, L090, L092, L094, L096, L099, L102, L103, L152, L160, L161, L164, L176, L179, L182, L188, L189, L192, L194, L210, L211, L212, L213, L214, L215, L216, L228, L229, L238, L255, L260, L262, L263, L264, L265, L267, L283, L292, L294									
WEOG	13	13	30	27	27	25	33	29	103	94
	L024, L026, L029, L031, L033, L034, L035, L036, L075, L098, L101, L104, L105, L107, L109, L110, L115, L117, L124, L125, L126, L128, L129, L130, L131, L132, L133, L134, L135, L136, L139, L141, L142, L143, L145, L146, L147, L150, L183, L184, L195, L208, L223, L224, L242, L246, L256, L257, L275, L276, L286, L287, L288, L290, L291, L298, L305									
Total	**103**	**82**	**105**	**89**	**176**	**133**	**148**	**116**	**532**	**420**

Table S 4: Laboratories that registered but did never deliver results

No regs.	Region	Lab
3x	Africa	L058
2x	GRULAC	L045
	Africa	L052
	Africa	L095
	Africa	L180
	GRULAC	L215
	CEE	L239
	Africa	L245

No regs.	Region	Lab
1x	Asia	L021
1x	Asia	L040
1x	GRULAC	L044
1x	GRULAC	L055
1x	Asia	L059
1x	Africa	L066
1x	GRULAC	L088
1x	Africa	L089
1x	GRULAC	L099
1x	WEOG	L109
1x	WEOG	L133
1x	Asia	L144
1x	WEOG	L150
1x	Asia	L159
1x	GRULAC	L160
1x	CEE	L162
1x	CEE	L165
1x	CEE	L168
1x	CEE	L172
1x	CEE	L174
1x	Africa	L177
1x	Asia	L178
1x	WEOG	L184
1x	GRULAC	L192
1x	Africa	L199

No regs.	Region	Lab
1x	Asia	L200
1x	GRULAC	L210
1x	GRULAC	L212
1x	GRULAC	L214
1x	GRULAC	L216
1x	Asia	L222
1x	Asia	L225
1x	GRULAC	L228
1x	Africa	L236
1x	CEE	L241
1x	Asia	L243
1x	WEOG	L246
1x	Asia	L247
1x	Asia	L248
1x	Asia	L249
1x	Asia	L251
1x	GRULAC	L260
1x	GRULAC	L263
1x	Africa	L273
1x	Africa	L274
1x	Africa	L281
1x	GRULAC	L283
1x	GRULAC	L292
1x	GRULAC	L294
1x	Africa	L295
1x	Asia	L299
1x	Asia	L300
1x	WEOG	L305

Figure S 1: Graphical sketch for each laboratory showing registration and results obtained within the round of the interlaboratory assessment

Table S 5: Summary of z-scores across four rounds of interlaboratory assessments by POP and type

	Test solution					Abiotic					Biota				
z-score No. of results	S 10126	Q 1550	U 3173	C 16	I 147	S 9413	Q 1466	U 4165	C 193	I 720	S 5653	Q 975	U 2967	C 375	I 636
aldrin	119	25	61	2	5	22	6	22	0	9	7	0	6	7	8
dieldrin	112	29	65	2	3	26	2	24	1	7	30	8	38	7	27
endrin	88	31	73	1	5	13	4	11	1	11	0	0	0	0	0
chlordane	440	66	178	0	9	183	15	105	34	55	153	35	98	30	47
DDT	730	167	403	4	17	504	105	485	21	148	299	65	315	32	99
heptachlor	289	59	191	2	17	71	13	57	7	22	61	9	58	5	19
mirex	85	10	27	0	0	45	6	26	2	5	39	9	29	2	11
toxaphene	87	6	13	0	0	0	0	0	0	0	4	2	3	0	0
a_HCH	80	23	56	1	3	48	8	53	2	27	13	1	7	2	4
b_HCH	65	20	65	0	6	35	8	43	4	32	52	8	40	14	37
lindane	86	15	61	0	7	50	2	35	3	27	0	0	0	0	0
chlordecone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
endosulfan	170	74	224	0	16	9	1	10	0	6	0	0	0	0	0
PCB6	1087	207	414	1	29	1088	253	801	22	72	748	181	692	13	96
HCB	109	18	44	2	4	78	19	69	3	12	53	15	49	3	13
PeCBz	47	3	14	0	3	33	6	22	2	8	20	2	11	9	6
HCBD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
PCDD	995	101	114	0	5	1349	210	310	6	25	556	91	188	61	58
PCDF	1406	136	189	0	7	1802	263	440	0	18	740	90	236	93	65
dl_PCB	1379	204	351	0	8	1665	223	721	31	144	1206	204	539	17	55
PBDE	547	88	182	1	0	598	75	254	17	35	386	84	168	26	44
PBDE_209	8	3	9	0	0	18	1	8	0	4	6	0	4	6	3
HBCD	125	5	15	0	0	16	2	1	0	4	45	0	16	4	0
HxBB	34	5	10	0	2	26	6	5	2	3	22	0	2	5	3
PFOS	149	22	20	0	1	189	22	91	4	6	231	17	57	12	2
PFOA	62	5	4	0	0	27	8	10	1	1	33	3	8	3	3
PFHxS	59	4	6	0	0	26	4	8	0	5	31	3	4	8	7
sum_drins	125	42	83	0	0	28	6	46	0	0	62	8	92	0	0
sum_HCHs	92	24	107	0	0	51	7	58	0	0	33	4	54	0	0
TEQ_DF	281	24	32	0	0	414	52	67	0	0	194	29	56	0	0
TEQ_PCB	224	36	54	0	0	296	35	144	0	0	182	34	64	0	0
TEQ_total	244	16	32	0	0	337	32	82	0	0	174	36	58	0	0
sum_PBDE	59	5	13	0	0	74	9	28	0	0	66	9	23	0	0
PFOSprec	98	12	19	0	0	40	6	19	2	0	12	1	0	0	1
nonSC_ PFAS	567	52	38	0	0	214	48	83	28	34	154	22	39	16	28
sum_ PFOSprec	18	5	1	0	0	10	2	7	0	0	0	0	0	0	0
sum_nonSC_ PFAS	60	8	5	0	0	28	7	20	0	0	41	5	13	0	0

Table S 6: Laboratories with most z-scores

Lab	S	Q	U	C	I	Total
L027	1171	86	55	18	5	1 335
L004	798	101	137	6	8	1 050
L011	753	68	197	8	22	1 048
L101	656	62	175	15	16	924
L117	711	52	99	32	22	916
L030	631	72	79	10	18	810
L126	572	63	129	12	27	803
L105	470	87	148	45	52	802
L024	630	94	41	16	16	797
L072	573	62	115	13	16	779
L132	424	136	150	14	11	735
L013	351	83	194	15	56	699
L145	527	58	64	16	13	678
L023	238	71	266	17	47	639
L001	520	46	59	4	3	632
L107	426	80	42	28	29	605
L005	460	55	67	2	8	592
L016	239	11	307	1		558
L153	406	39	112			557
L025	502	29	22			553
L190	347	42	136	4	15	544
L128	339	63	80	4	4	490
L022	293	76	115	1		485
L015	280	92	108	1	1	482
L002	390	36	50	3	2	481
L008	383	33	63			479
L053	185	99	182			466
L124	376	30	32	12	4	454
L173	279	63	108		2	452
L125	384	29	9	5	1	428
L115	233	92	91	6	4	426
L037	349	32	27		2	410
L007	227	53	126	1		407
L137	297	28	63	2	7	397
L003	253	19	107			379

Lab	S	Q	U	C	I	Total
L034	316	21	18	2		357
L060	176	54	105	5	16	356
L073	172	54	122	3	3	354
L242	300	11	30	4	1	346
L134	215	22	56	17	28	338
L035	185	52	97			334
L019	83	36	199	5	8	331
L031	190	44	60	7	5	306
L050	97	27	121	4	45	294
L091	48	20	148	4	68	288
L104	184	45	56		1	286
L148	179	32	69			280
L012	224	21	29			274
L102	84	30	124	5	31	274
L065	133	15	88	5	25	266
L062	14	13	147	10	79	263
L112	160	31	60			251
L017	198	15	21	6	9	249
L135	118	16	73	10	31	248
L195	124	26	84		2	236
L156	145	37	40	1	9	232
L041	87	23	121			231
L029	181	16	26			223
L094	183	5	9	8	8	213
L163	13	6	177	1	16	213
L147	129	36	37	1		203

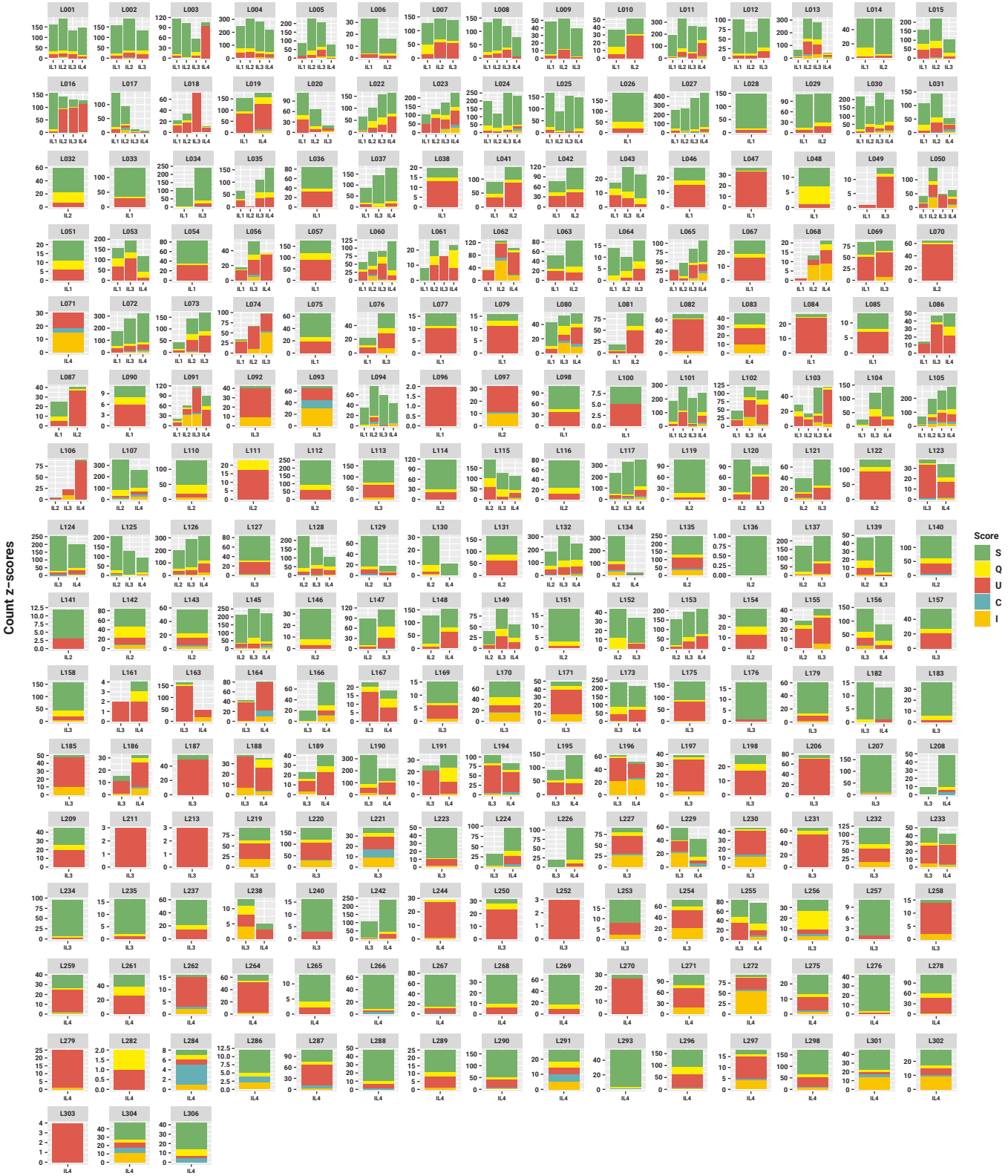


Figure S 2: Overview of the participation and the performance of labortories in the various rounds

UNITED NATIONS ENVIRONMENT PROGRAMME
United Nations Avenue, Gigiri
P O Box 30552, 00100, Nairobi, Kenya
Tel +254 720 200200
unep-info@un.org
www.unep.org