

First inter-laboratory comparison exercise for the determination of anticancer drugs in aqueous samples

Ester Heath^{1,2} · Marjeta Česen^{1,2} · Noelia Negreira³ · Miren Lopez de Alda³ ·
Laura Ferrando-Climent⁴ · Lucie Blahova⁵ · Tung Viet Nguyen⁶ ·
Mohamed Adahchour⁷ · Achim Ruebel⁸ · Neville Llewellyn⁹ ·
Janez Ščančar^{1,2} · Srdjan Novaković¹⁰ · Vesna Mislej¹¹ ·
Marjeta Stražar¹² · Damià Barceló^{3,4} · Tina Kosjek^{1,2}

Received: 27 April 2015 / Accepted: 29 June 2015 / Published online: 15 July 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract The results of an inter-laboratory comparison exercise to determine cytostatic anticancer drug residues in surface water, hospital wastewater and wastewater treatment plant effluent are reported. To obtain a critical number of participants, an invitation was sent out to potential laboratories identified to have the necessary knowledge and instrumentation. Nine laboratories worldwide confirmed their participation in the exercise. The compounds selected (based on the extent of use and laboratories capabilities) included cyclophosphamide, ifosfamide, 5-fluorouracil, gemcitabine, etoposide, methotrexate and cisplatin. Samples of spiked waste (hospital and wastewater treatment plant effluent) and surface water, and additional non-spiked hospital wastewater, were prepared by the organising laboratory (Jožef Stefan Institute) and sent out to each participant partner for analysis. All analytical methods

included solid phase extraction (SPE) and the use of surrogate/internal standards for quantification. Chemical analysis was performed using either liquid or gas chromatography mass (MS) or tandem mass (MS/MS) spectrometry. Cisplatin was determined using inductively coupled plasma mass spectrometry (ICP-MS). A required minimum contribution of five laboratories meant that only cyclophosphamide, ifosfamide, methotrexate and etoposide could be included in the statistical evaluation. *z*-score and *Q* test revealed 3 and 4 outliers using classical and robust approach, respectively. The smallest absolute differences between the spiked values and the measured values were observed in the surface water matrix. The highest within-laboratory repeatability was observed for methotrexate in all three matrices ($CV \leq 12\%$). Overall, inter-laboratory reproducibility was poor for all compounds and matrices

Responsible editor: Philippe Garrigues

Electronic supplementary material The online version of this article (doi:10.1007/s11356-015-4982-9) contains supplementary material, which is available to authorized users.

✉ Ester Heath
ester.heath@ijs.si

¹ Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

² International Postgraduate School Jožef Stefan, Jamova 39, Ljubljana, Slovenia

³ Water and Soil Quality Research Group, Institute of Environmental Assessment and Water Research (IDAEA-CSIC), Jordi Girona 18, 08034 Barcelona, Spain

⁴ Catalan Institute for Water Research, Emili Grahit 101, Edifici H2O, Parc Científic i Tecnològic, 17003 Girona, Spain

⁵ Faculty of Science, RECETOX, Masaryk University, Kamenice 5 Bmo, Czech Republic

⁶ National University of Singapore, 21 Lower Kent Ridge Rd, Singapore, Singapore

⁷ Eurofins Omegam, H.J.E. Wenckebachweg 120, 1114 AD Amsterdam-Duivendrecht, The Netherlands

⁸ IWW Water Centre, Moritzstrasse 26, 45476 Muelheim, Germany

⁹ CEH Lancaster, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, UK

¹⁰ Institute of Oncology, Zaloška cesta 2, 1000 Ljubljana, Slovenia

¹¹ Central Wastewater Treatment Plant Ljubljana VO-KA, Cesta v prod 100, 1000 Ljubljana, Slovenia

¹² Wastewater Treatment Plant Domžale-Kamnik, Študljanska cesta 91, 1230 Domžale, Slovenia

(CV 27–143 %) with the only exception being methotrexate measured in the spiked hospital wastewater (CV=8 %). Random and total errors were identified by means of Youden plots.

Keywords Inter-laboratory study · Cytostatic · Pharmaceutical · Surface water · Wastewater · Hospital effluent

Introduction

The World Cancer Report (2014) published by the World Health Organization (WHO) states that cancer cases are expected to surge 57 % worldwide over the next 20 years, a statistic that is, in the main, attributable to aging global populations (World Cancer Report 2014, IARC, WHO). This is also leading to an increase in the use of chemotherapeutic medicines (CytoThreat Project 2014), which now represents the third largest source of revenue within the pharma industry (Reuters 2008 and Chemoth 2013). These drugs are typically administered to outpatients in oncological units (75 %) using ambulatory infusion, albeit a more recent trend is towards oral administration at home (Kosjek and Heath 2011). After receiving therapy, patients excrete anticancer drugs in both the parent and metabolised forms (Cytotoxic Safety 2013a, b). Excretion rates are drug dependent but, in some cases, can take more than a week, and the compounds can remain in an active stable form for significant periods of time. Illustrative examples are cyclophosphamide and ifosfamide, common anticancer drugs classified as alkylating agents. These two agents under ideal laboratory conditions can remain intact for up to 800 days in the case of cyclophosphamide and 10 to 12 years for ifosfamide, implying that these chemicals could survive under environmental conditions for long periods (Kosjek and Heath 2011). Once excreted from the body, they end up in the sewerage system eventually to arrive at a municipal wastewater treatment plant (WWTP), where studies have demonstrated that they are not completely removed (Kosjek and Heath 2011). Instead, they pass through the treatment process into surface water raising concerns over their presence in the environment and in potable water supplies (Cytotoxic Safety 2013a, b; Crauste-Manciet et al. 2005). This means that although beneficial for the cancer patient, as a consequence of their cytotoxic, genotoxic, mutagenic and/or teratogenic properties cytostatic drug residues may pose a threat to non-target organisms and human health (Cytotoxic Safety 2013a, b).

Until recently, scientists did not have the technology nor the know-how to measure environmental levels of anticancer drugs and only a limited number of laboratories worldwide are

analysing them in environmental and wastewater samples (Česen et al. 2015; Ferrando-Climent et al. 2014, 2013; Kosjek and Heath 2011; Kosjek et al. 2015, 2013; Llewellyn et al. 2011; Negreira et al. 2014a, 2013a, b; Parrella et al. 2014). One reason is their low environmental levels ($\leq \text{ng L}^{-1}$), which require the very latest in modern analytical instrumentation, and possibly their toxicity makes laboratories reluctant to analyse these compounds since additional safety protocols and waste collection must be in place. Fortunately, studies are now investigating their presence in environmental and wastewater samples (Česen et al. 2015; Kosjek and Heath 2011; Kosjek et al. 2015, 2013; Negreira et al. 2014a, 2013a, b), and the risks they pose to humans and aquatic organisms have been recently addressed within the EU FP7 CytoThreat project (2014).

In the absence of certified reference materials, inter-laboratory comparisons are necessary to verify analytical results, if laboratories are to have confidence in their analytical abilities. A number of inter-laboratory studies addressing emerging contaminants in environmental samples have been carried out (Farré et al. 2008; Heath et al. 2010a, b; Hund et al. 2000; Jones et al. 1994; Van den Bossche et al. 2010; Vander Heyden and Smeyers-Verbeke 2007), of which only a few address veterinary drug residues (Vander Vander Heyden et al. 1999), nonsteroidal anti-inflammatory drugs and hormones (Farré et al. 2008; Heath et al. 2010a, b) and antibiotics (Dehouck et al. 2003; Van den Bossche et al. 2010). To our knowledge, none included cytostatic drugs. This study was set out to find those laboratories analysing cytostatic drug residues at trace levels in environmental and wastewater samples and to include them in the first inter-laboratory exercise on cytostatic drugs in water to facilitate and promote knowledge exchange between them, evaluate their performance in the analysis of these compounds when using their own in-house analytical methods, and elucidate potential bias and sources of error.

Experimental

Experimental design, sample collection and handling

An invitation to participate in the exercise was offered to those laboratories identified as having the necessary knowledge and instrumentation to determine anticancer residues in environmental and wastewater samples. Participating laboratories were from the following countries: Australia, Czech Republic, Germany, The Netherlands, Singapore, Slovenia, Spain and UK. Each participating laboratory was assigned its own inter-laboratory code (L1 to L9).

Four different samples were included in the exercise: a surface river water spiked with the analytes at a predefined concentration (A), a hospital wastewater (B), the same

hospital wastewater spiked with the target compounds at a known concentration (C) and a spiked municipal WWTP effluent (D). Prior to the exercise, the river water and the WWTP effluent samples had been analysed in the laboratory and found not to contain tested drug residues. These samples were used in the study after spiking them with appropriate environmentally and wastewater relevant concentrations of the target analytes. Analysis of hospital wastewater revealed the presence of several cytostatic residues including CP, IF, 5-FU, GEM, ETO and MTX, but not all those included in this study (cis-Pt). Both samples as collected from the hospital and spiked with the target compounds were included in the trial.

All three matrix samples (the surface water, the WWTP effluent and the hospital effluent) were collected by grab sampling and are actually interrelated since the surface water sample was collected from one of the largest rivers in Slovenia approximately 100 m downstream from the WWTP effluent outflow, and the WWTP receives the effluent wastewater from the hospital. All samples were collected in clean polyethylene containers and transferred on ice immediately to the laboratory, where they were filtered (0.5- μm glass fibre filters) and homogenised.

All the samples, with the exception of sample B, were spiked (Table 1) with the target compounds: cyclophosphamide (CP), ifosfamide (IF), 5-fluorouracil (5-FU), gemcitabine (GEM), etoposide (ETO), methotrexate (MTX) and cisplatin (Cis-Pt) at environmental and wastewater levels (Kosjek and Heath 2011).

Samples were stored in polyethylene containers at $-80\text{ }^{\circ}\text{C}$. Frozen samples (0.8 L) were shipped on dry ice to each of the participating laboratories. On receipt, the time of arrival and sample condition were recorded. Each laboratory was also asked to process their samples, i.e. filtration and extraction, within 1 week. Finally, the participating laboratories were asked to perform at least two independent analyses per sample. The stability of the compounds CP, IF, 5-FU, GEM, ETO and MTX in HPLC water and in wastewater samples had been studied previously and found to be acceptable under the inter-laboratory study conditions (Ferrando-Climent et al. 2013; Negreira et al. 2014b).

Chemicals

The target compounds CP (CAS 50-18-0), IF (CAS 3778-73-2), 5-FU (CAS 51-21-8) and GEM (CAS 95058-81-4) were

obtained from Sigma-Aldrich (Steinheim, Germany), MTX (CAS 59-05-2) was obtained from TOCRIS Biosciences (Ellisville, USA) and ETO (CAS 33419-42-0) was obtained from Santa Cruz Biotechnology (Heidelberg, Germany). Dimethylsulfoxide (DMSO) was obtained from Sigma-Aldrich (St. Luis, USA). Fresh standard solutions of a mixture of the cytostatic compounds were prepared on a weight basis in DMSO.

Analytical methods

Chemical analysis was performed using previously developed analytical methods based on either liquid or gas chromatography (LC or GC) coupled to mass (MS) and most frequently tandem mass (MS/MS) spectrometry (Table 2) (Česen et al. 2015; Ferrando-Climent et al. 2013; Kosjek et al. 2013; Llewellyn et al. 2011; Negreira et al. 2013a, b, 2014b; Odraska et al. 2013; Yin et al. 2010). All analytical methods, with the exception of that based on ICP-MS for Pt determination, included a solid phase extraction (SPE) step, performed either on-line or off-line for sample preconcentration. Internal standards (IS) were used for quantification (Table 2). Laboratory 9, which analysed samples by GC-MS and GC-MS/MS, also included a derivatization step, using trifluoroacetic acid to derivatise CP and IF and N-(tert-butyl dimethylsilyl)-N-methyltrifluoroacetamide (MTBSTFA) to derivatise 5-FU.

Statistical parameters

Sample homogeneity was tested for each batch of samples using the chi-square test (Eq. 1):

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (1)$$

where O_i represents the mean concentration of two parallels of each sample and E_i is the mean concentration of each batch containing 10 samples (Heath et al. 2010b). The null hypothesis (H_0) states that homogeneity of samples is achieved, while the alternative hypothesis (H_1) says that homogeneity of samples is not achieved. If χ^2_{tab} ($\alpha=0.05$) is greater than χ^2_{exp} , then the H_0 is not rejected. To perform the chi-square test, ten random subsamples of each sample (A, C and D) were collected and each sample analysed in duplicate.

Table 1 Spiked matrices and concentrations

Matrix/compound	CP (ng L ⁻¹)	IF (ng L ⁻¹)	5FU (ng L ⁻¹)	GEM (ng L ⁻¹)	ETO (ng L ⁻¹)	MTX (ng L ⁻¹)	Cis-Pt (ng L ⁻¹)
A: Surface river water	53	33	43	21	52	31	17
C: Hospital wastewater	5239	394	454	397	3420	1141	403
D: WWTP effluent	23	55	66	786	127	432	35

CP cyclophosphamide, IF ifosfamide, 5FU 5-fluorouracil, GEM gemcitabine, ETO etoposide, MTX methotrexate, Cis-Pt cisplatin

Table 2 General information on the applied analytical methods including analytical technique, sample volume, sample pre-treatment and the internal standards used, detection conditions and limits of

detection (LOD). A detailed description of the method is in the supplementary material (SM)—Tables S1–S6 and S8–S9 (L7 could not analyse the samples due to their destruction at customs)

Lab code	Instrumentation	Sample volume	SPE	Internal standard (addition before or after extraction)	Detection	LOD ^a
1	LC-MS/MS	5 mL	On-line SPE with PLRP-s (10×2 mm)	IF-d ₄ (before) CP-d ₄ (before) ETO-d ₃ (before) GEM- ¹³ C, ¹⁵ N ₂ (before) MTX-methyl-d ₃ (before)	QTRAP (QqLIT) MS/MS ESI+	IF 0.2 ng L ⁻¹ CP 0.1 ng L ⁻¹ ETO 3.0 ng L ⁻¹ GEM 0.3 ng L ⁻¹ MTX 0.1 ng L ⁻¹
2	LC-MS/MS	300 mL	Strata X (200 mg/6 mL)	CP-d ₄ (before)	QqQ; MS/MS mode ESI+	CP 0.233 ng L ⁻¹
3	LC-MS/MS	250 mL	Double SPE on WAX (6 cc, 150 mg, 30 μm) and HLB (6 cc, 200 mg, 30 μm)	Caffeine d ₉ (before) NP (after)	QTRAP ESI+	CP 1.5 ng L ⁻¹ ETO 2 ng L ⁻¹ IF 1.5 ng L ⁻¹ GEM 1.5 ng L ⁻¹
4	UPLC-MS/MS	50 mL	HLB (200 mg/6 mL)	CP-d ₄ (after) ETO-d ₃ (after) MTX-d ₃ (after)	QTRAP (QqLIT) MS/MS ESI+	CP 1.1 ng L ⁻¹ ETP 6.4 ng L ⁻¹ IF 1.7 ng L ⁻¹ MTX 1.8 ng L ⁻¹
5	HPLC-ICP-MS	15 mL	–	External calibration	ICP-MS	cis-Pt 300 ng L ⁻¹
6	LC-MS/MS	4 mL	on-line SPE with HySphere Resin GP (10×2 mm)	Isoproturon-d ₆ (before) Tribromophenol (after)	QqQ; MS/MS mode (ESI+)	5-FU 100 ng L ⁻¹ CP 1.2 ng L ⁻¹ ETO 3 ng L ⁻¹ IF 1.3 ng L ⁻¹ MTX 50 ng L ⁻¹
8	LC-MS/MS	500 mL	Strata X (500 mg/6 mL) Florisil (200 mg/3 mL)	CP-d ₄ (before)	QqQ; MS/MS mode (ESI+)	CP 0.03 ng L ⁻¹ IF 0.05 ng L ⁻¹
9	GC-MS/MS	100 mL (5-FU)	Isolute ENV+ (500 mg sample A and 1 g for samples B, C, D/6 mL)	5FU-d ₆ and 5-chlorouracil (before)	Q; MS-mode (EI): SIM	5FU 0.16 ng L ⁻¹
	GC-MS	100/200/400 mL (CP, IF)	Oasis HLB (60 mg/3 mL)	CP-d ₆ (for CP and IF, before)	IT; MS-MS mode (EI)	CP 0.6 ng L ⁻¹ IF 0.3 ng L ⁻¹
	ICP-MS	100 mL (cis-Pt)	–	–	ID-ICP-MS	Cis-Pt 0.6 μg Pt L ⁻¹

EI electron impact, ESI electrospray, IT ion trap, NP nonylphenol, QqLIT quadrupole linear ion trap, QqQ triple quadrupole, SIM selected ion monitoring, ID isotope dilution

^aLowest LOD determined (see Supplementary material Table S1–S9)

Outlier detection was performed by calculating z-score (Heath et al. 2010a, b) values according to Eq. 2,

$$z = \frac{\chi_{lab} - \chi_0}{\sigma_0} \tag{2}$$

where χ_{lab} is the laboratory mean (classical approach) or median (robust approach), χ_0 is the known spiked concentration or, if unknown, as in sample B, is the average concentration measured by the participating laboratories and σ_0 is the standard deviation. Samples with z-values higher than 3 were excluded from further statistical analysis. For suspected outliers, i.e. those whose z-score values were between $2.0 \leq |z| \leq 3.0$, a

further Dixon test or Q test was applied ($\alpha=0.05$) (Heath et al. 2010b). The equation for the Q test (Heath et al. 2010b) is as follows (Eq. 3):

$$Q = \frac{\text{gap}}{\text{range}} \tag{3}$$

Gap is the absolute difference between the suspected outlier and its closest value when arranged in increasing order of concentration, and the range is the difference between highest and lowest value. When Q_{exp} is greater than Q_{tab} (a reference

value corresponding to the sample size and confidence level or $\alpha=0.05$), then this value is referred as an outlier.

The following statistical parameters were calculated for each series of samples (A, B, C and D) and each compound:

- Mean and median values;
- 95 % confidence intervals;
- Variances (σ^2);
- Standard deviations (σ);
- Relative standard deviation (RSD);
- Standard errors of mean (SEM);
- Minimum and maximum values and first and third quartiles (25P and 75P);
- Number of outliers, upper and lower warning limits (UWL and LWL); and
- Repeatability (r) and reproducibility (R), where both r and R were determined in terms of relative standard deviation or coefficient of variation (CV) and expressed as percentages (%).

The equations used to calculate the UWLs and LWLs are as follows (Eqs. 4 and 5):

$$\text{UWL} = (\bar{x} + 2\sigma) \quad (4)$$

$$\text{LWL} = (\bar{x} - 2\sigma) \quad (5)$$

The repeatability for each sample (A, B, C and D) and selected compound was individually determined for each participating laboratory as the CV corresponding to the ratio between the standard deviation of the laboratory's measurement and its mean concentration value (ISO TC 69/SC 6 N 2011; Eq. 6):

$$\text{CV}(\%) = 100 \times \frac{\sigma_{\text{lab}}}{\chi_{\text{lab}}} \quad (6)$$

The reproducibility was determined as the CV, where the ratio between the standard deviation of all laboratories and the average value of all laboratories were calculated for each compound and sample (ISO TC 69/SC 6 N 2011). All statistical evaluation was performed using MedCalc Software and Excel 2010.

Results and discussion

The stability of CP, IF, 5-FU, GEM, ETO and MTX in HPLC grade water and wastewater samples has been studied by both Ferrando-Climent et al. (2013) and Negreira et al. (2014b)). Their results reveal that these compounds are stable for at least 1 month at -20°C even in wastewater. Furthermore, the authors

(2013) found that CP, IF and MTX were stable for 3 months when stored at -20°C on SPE Oasis HLB cartridges. Despite this, we requested that the participating laboratories extract their samples within 1 week. Eight laboratories managed to extract the samples within 1–3 days, while laboratory 9 reported a 5-month delay when analysing 5-FU. Unfortunately, samples shipped overseas were destroyed at customs. Sample homogeneity was evaluated using the chi-square test (Eq. 1). The experimental values of χ^2_{exp} were lower than χ^2_{tab} ($\alpha=0.05$) for each sample type and selected compounds (CP, IF and MTX, (Hund et al. 2000); ISO 13528 2009) proving that the samples were homogeneous (data not shown).

In total, 266 data were received of which 219 were used for statistical evaluation. The low number of laboratories analysing certain compounds (e.g. only two laboratories submitted results for 5-FU and GEM, while for cis-Pt, only one laboratory submitted results that were above the LOD in all matrices) meant that it was not possible to use all the data. Table 3 lists the concentrations reported and the average and standard deviation values for GEM and 5-FU, while the results for CP, IF, MTX and ETO together with their corresponding statistical evaluation by classical and robust approaches are presented in SM Table S10 and Table S11, respectively. To take into account the most probable scenario, when the concentrations of analytes of interest were reported to be below LOQ, a value equivalent to half the reported LOQ was used for the evaluation (European Food Safety Authority 2010).

Laboratory performance regarding GEM (L1 and L3) and 5-FU (L6 and L9) are based on two submitted data sets (Table 3). In the case of GEM, the results showed a fairly good performance by L1 for all 4 samples analysed and by L3 for the surface water sample. However, in the case of the wastewater samples (B, C, D), the values reported by L3 were lower and inconsistent with the spiked values indicating that matrix effects were not sufficiently taken into account with the IS used. In the case of 5-FU, the low number of laboratories providing data (only L6 and L9) may be due to the complex nature (number of steps) of the analytical procedure necessary for determining this compound at trace levels, and because this compound, in contrast to most cytostatics, is more easily ionised in the negative than that in the positive ion mode (Kosjek et al. 2013). This is also the most likely reason behind the high LOD obtained by L6 for 5-FU (100 ng L^{-1} , in the ESI+ mode), which prevented it from being detected at the levels spiked in the surface water (sample A, 43 ng L^{-1}) and in the WWTP effluent (sample D, 66 ng L^{-1}). For L6, the value reported for the spiked hospital wastewater (583 ng L^{-1}) was higher than the spiked concentration (454 ng L^{-1}), but not consistent with the value measured in the non-spiked sample (420 ng L^{-1}) to account for the sum of both concentrations ($420\text{ ng L}^{-1} + 454\text{ ng L}^{-1} = 874\text{ ng L}^{-1}$). Interpretation of the possible reasons in this case is difficult. In the case of L9, however, the values measured in the four samples analysed by GC-MS/MS were half to three quarters of the

Table 3 Data reported for GEM and 5-FU by two participating laboratories: raw data from three independent replicate analyses, calculated average and standard deviation and spiked concentration in each sample type

Laboratory code	Sample					Average	SD	Spiked conc. (ng L ⁻¹)
GEM								
L1	A	19.4	23.4	29.1	24.0	4.9	21	
	B	13.6	13.1	–	13.4	0.4	0	
	C	402	434	416	417	16	397	
	D	704	698	600	667	58	786	
L3	A	17	19	23	19.7	3.1	21	
	B	231	294	293	273	36	0	
	C	191	256	295	247	53	397	
	D	170	280	348	266	90	786	
5-FU								
L6	A	150 ^a	150 ^a	150 ^a	150 ^a	–	43	
	B	440	400	420	420	20	0	
	C	610	560	580	583	25	454	
	D	150 ^a	150 ^a	150 ^a	150 ^a	–	66	
L9	A	19	19	19	19	0	43	
	B	15	11	8.3	12	3	0	
	C	315	364	247	309	59	454	
	D	39	41	38	39	5	66	

^a Values reported below LOQ; therefore, half LOQ taken into account for evaluation

spiked value, a discrepancy explained on the basis of the stability results reported by Negreira et al. (2014b)), who found that for wastewater samples stored for 3 months at –20 °C, 25 % of the spiked 5-FU was degraded. In the case of L9, storage time was 5 months.

Classical and robust statistical approaches applied to the data reported for CP, IF, MTX and ETO did not show relevant differences in the number of outliers, which occur only in sample B (non-spiked hospital wastewater): three outliers (for CP) using the classical approach and four outliers (three

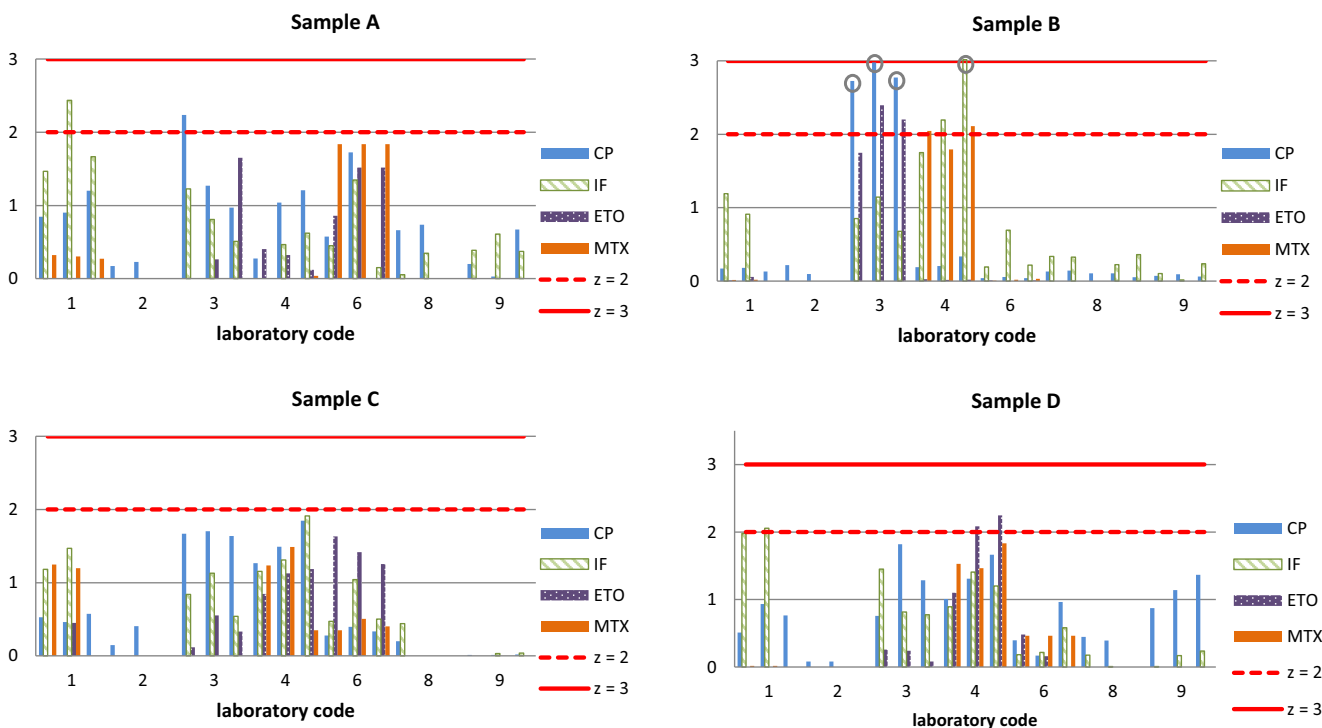


Fig. 1 Graphical presentation of z-score values for each compound in the various samples using a robust approach

for CP and one for IF) using the robust approach (SM—Table S12). In total, out of the nine laboratories, L3 and L4 together produced four identified outliers using the robust approach (Fig. 1), while L3 produced three outliers determined using the classical approach (SM—Figure S1). The group L3 experienced difficulties in determining CP, while L4 found determining IF problematic when analysing hospital wastewater. L4, despite extracting only 50 mL of sample, reported acceptable LOD for all tested compounds (low ng L^{-1} , Table 2). The poor analytical performance in determining IF is thought to be a result of not using a deuterated analogue of IF as the internal standard. Similarly, L3 who extracted a larger sample volume (250 mL) compared to L4 (50 mL), also applied structurally different compounds as internal standards and could have improved their method performance by using isotopically labelled analogues as internal standards.

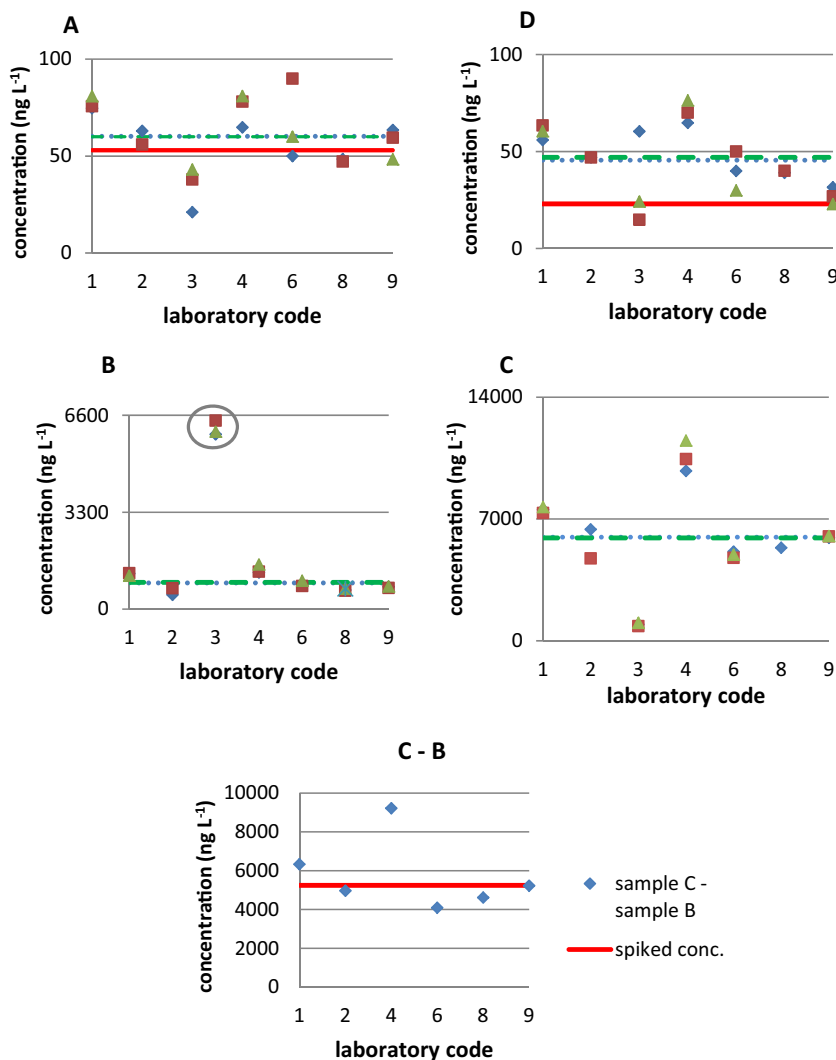
The identification of outliers only in sample B can be attributed to the complexity of the matrix (hospital wastewater) and the comparatively lower concentrations of the analytes

when compared to sample C (spiked hospital wastewater), which produced no outliers.

Tables S10 and S11 (SM) show the statistical data determined once the outliers have been removed using both classical and robust approaches, respectively. Again, the only difference between the two data sets occurs in the case of IF in sample B. The data also show an agreement between the values for CP, IF and ETO in samples B and C, where $[C]=[B]+[\text{spiked}]$. In all cases, the values obtained in sample C are slightly lower than those calculated on the basis of the concentration determined in sample B. In the case of MTX, the data values differed significantly from the spiked values in all three matrices, especially in the hospital wastewater, where the measured concentration of MTX was higher in the non-spiked (Sample B; 5832 ng L^{-1}) than that in the spiked sample (sample C; 2271 ng L^{-1}).

Figures 2, 3, 4 and 5 show the concentrations of CP, IF, ETO and MTX, respectively, reported by the various laboratories for all four analysed samples (different characters

Fig. 2 Reported CP concentrations (different dots) in relation to spiked concentrations (red line) and inter-lab exercise mean (dotted blue line) and median (dashed green line) values calculated using robust approach



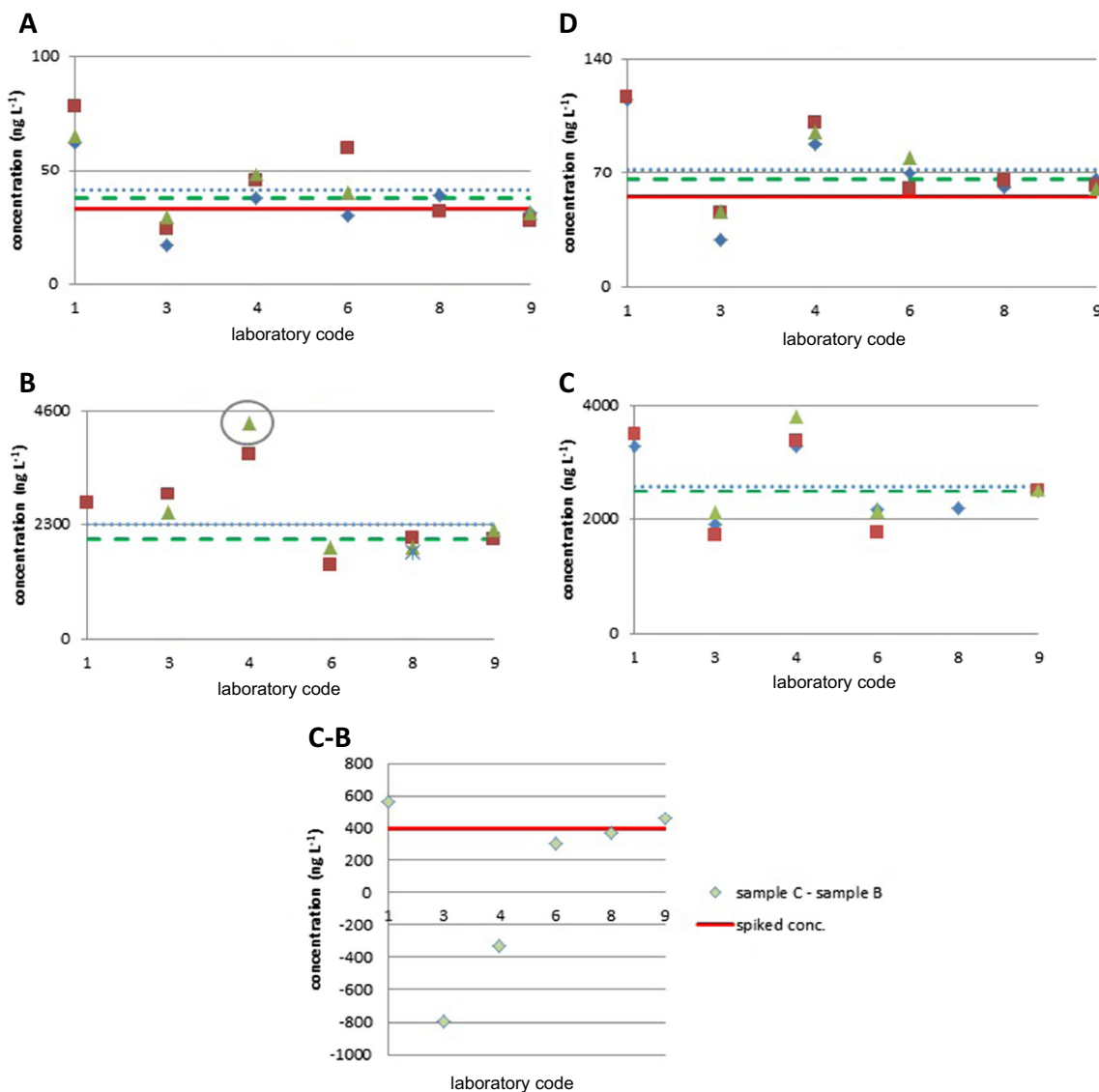


Fig. 3 Reported IF concentrations (*different dots*) in relation to spiked concentrations (*red line*) and inter-lab exercise mean (*dotted blue line*)

and median (*dashed green line*) values by calculated using robust approach

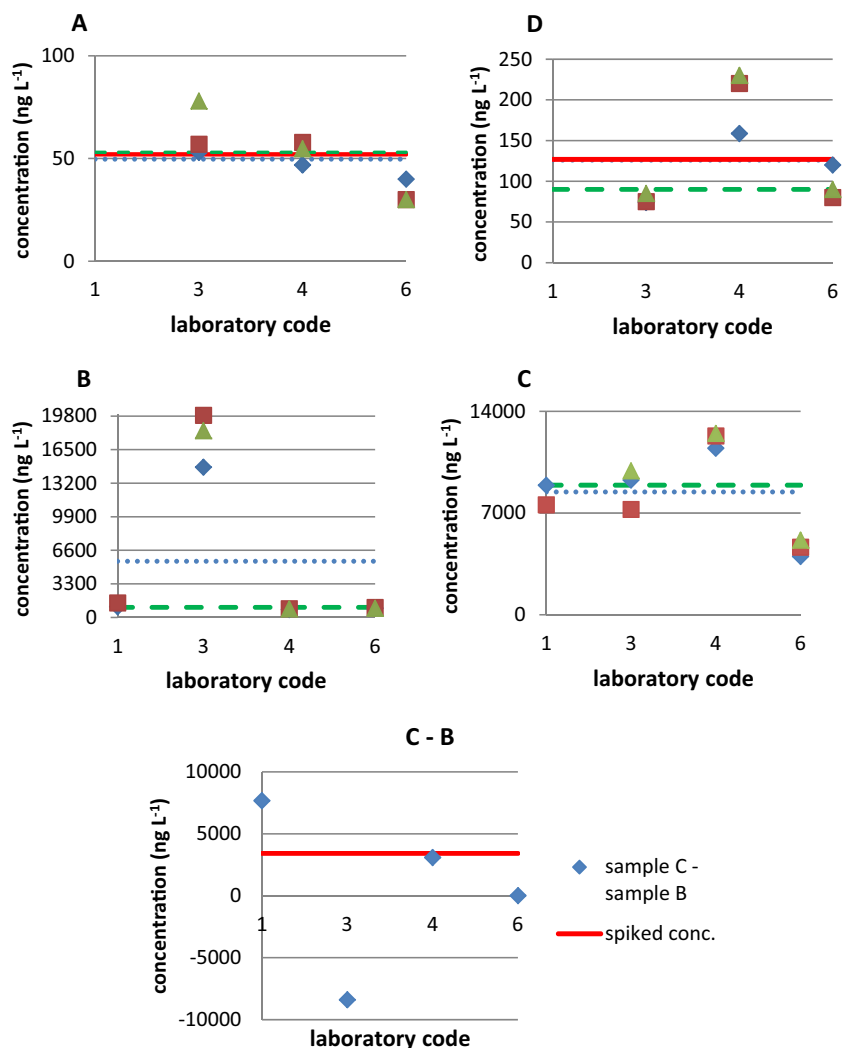
representing analytical replicates), the spiked concentrations (red line) and the calculated mean (dotted blue line) and median (dashed green line) values using robust approach. The data for IF using a classical approach is presented in SM—Figure S2 showing the only difference between the two data sets. Outliers (circled) are not included in the calculation of the mean and median values. The final graph for each figure represents the difference between spiked hospital wastewater and non-spiked hospital wastewater (C-B), where the red line represents the spiked value and dots are the average values calculated for each laboratory. Mean and median values of all the laboratories were relatively close to the spiked concentrations in samples A and D for all four compounds. Only in a few cases were negative values observed in the calculation of the difference between samples C and B when determining IF (L3

and L4), ETO (L3) and MTX (L4), i.e. compounds spiked in the hospital water at concentrations comparatively lower than those naturally present in the sample. Negative values were not recorded for CP (and also 5-FU and GEM), where spiked concentrations were higher than naturally occurring levels.

Method repeatability (*r*), expressed as CV, for each sample (A, B, C and D) and compound of interest is presented in SM—Figure S3. The lowest CVs were observed in the case of MTX ($\leq 12\%$ for L4, sample C) and the highest in the case of CP ($\leq 72\%$ for L3, sample D). Empty cells correspond to analyses where at least one of the replicate results was below the LOD or not reported.

Figure 6 shows a plot of the average CV (average of the various laboratory method repeatabilities) vs. the calculated mean concentration (average of the mean concentrations

Fig. 4 Reported ETO concentrations (different dots) in relation to spiked concentrations (red line) and inter-lab exercise mean (dotted blue line) and median (dashed green line) values calculated using robust approach



reported by the various laboratories) for CP, IF, ETO and MTX, i.e. the compounds determined by most laboratories for the four tested matrices. The CVs were expected to be highest for the lowest concentrations, i.e. in surface water (Heath et al. 2010b), but in reality, the CVs for CP are constant over the concentration range. This is a likely result of using a deuterated CP compound as internal standard (Table 2).

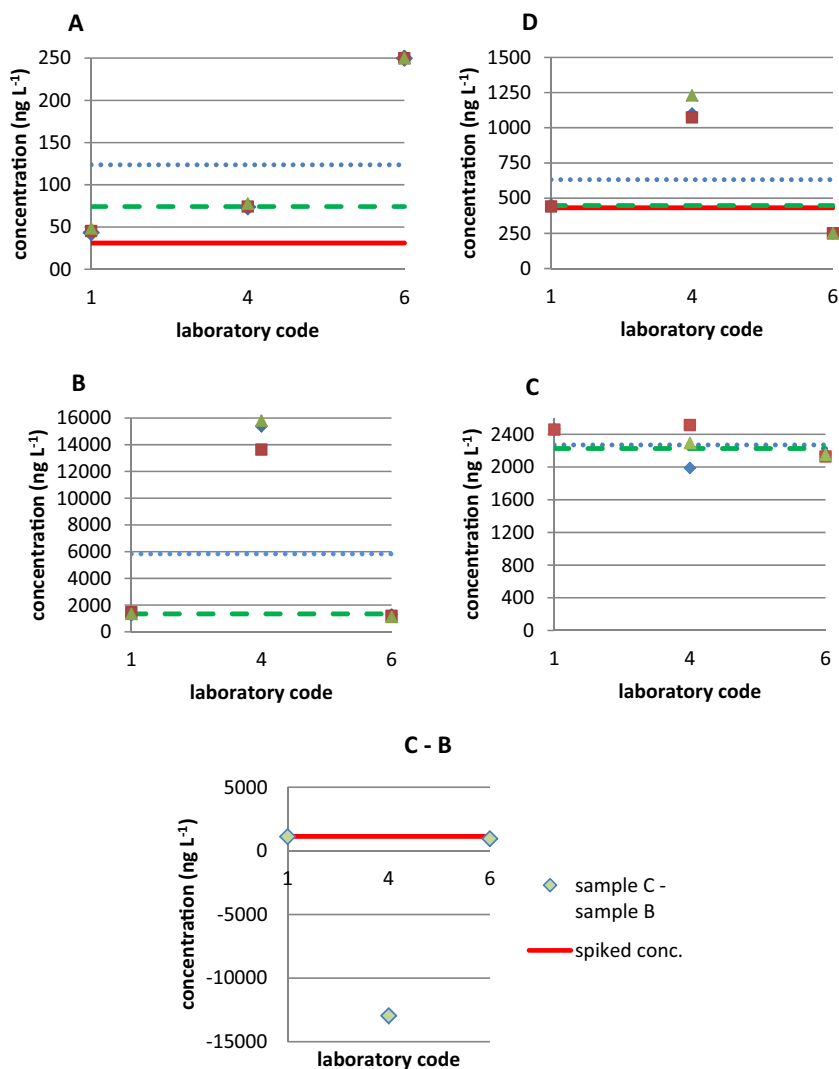
The volume of sample extracted varies between laboratories. Laboratories L2, L3, L8 and L9 extracted 100–500 mL of sample; L4 extracted 50 mL of sample, while L5 extracted 15 mL of sample (Table 2). Both L1 and L6, which perform on-line SPE, extracted the lowest amount of sample: 5 and 4 mL, respectively. A comparison of laboratory performance reveals outliers in the case of L3 and L4 indicating that method performance cannot be directly linked to sample extraction volume.

A comparison of the CV values obtained for the different compounds, matrices and laboratories shows similar CVs between samples B and C, both corresponding to hospital wastewater, even if the CV values were expected to be lower in the

spiked sample (C, SM—Figure S4). Likewise, even though high CVs were expected in the WWTP effluent (sample D) because of matrix complexity, no relevant difference in the CVs was observed between the surface water and the treated wastewater (samples A and D). The only exception in this case was the distinctly higher CV obtained by L3 when determining CP in sample D.

The reproducibility (R), i.e. the variability in the results reported by the various laboratories, expressed also as CV, for each sample (A, B, C and D) and the compounds CP, IF, ETO and MTX was also calculated (SM—Figure S5). Variability is mostly within 40 %, with the only exceptions being ETO and MTX in sample B. Since variability for these compounds is high in sample B, but not in sample C, it cannot be attributed to distinct matrix effects, since the matrix is the same for both B and C. This high variability derives from the high concentrations reported for ETO and MTX by L3 and L4, respectively, in sample B ($>14 \mu\text{g L}^{-1}$), concentration values that are difficult to interpret from an analytical perspective, especially considering that L4 used MTX- d_3 as an IS, unless the difference is down to human

Fig. 5 Reported MTX concentrations (different dots) in relation to spiked concentrations (red line) and inter-lab exercise mean (dotted blue line) and median (dashed green line) values calculated using robust approach



error. The inter-laboratory data were also analysed graphically (all laboratories having analysed two samples) using a Youden plot (MedCalc Software 2015), which visualises within-laboratory and between-laboratory variability. In the original Youden plot (Youden 1959), axes are drawn to the same scale

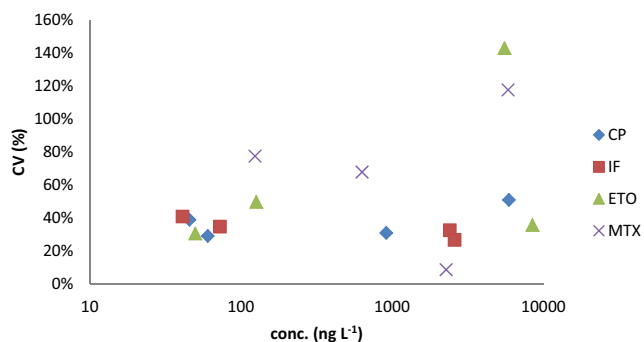


Fig. 6 Mean repeatability (CV) vs. concentration (ng L⁻¹) for various tested compounds

and the two samples must be similar and close in magnitude. Results of individual laboratories are defined by the response variables on the horizontal and vertical axis and are presented as points in the plot. The two median lines, one parallel to x and the other parallel to y axis, are drawn dividing results evenly. Around their intersection, a Manhattan circle is drawn including 95 % of laboratories and a 45° reference line through the Manhattan median. In the Youden plot, the points that lie near the 45° reference line but far from the Manhattan median indicate large systematic error, points far from the 45° line indicate large random error and points outside the circle indicate a large total error.

Youden plots were calculated for CP and IF in samples A and D due to similar spiked concentrations although the matrix complexity differs greatly. The plots reveal the nature of errors for each laboratory. Points outside the circle indicate large total error, which was observed for L1, L3 and L4 in case of both compounds. Furthermore, random error was observed in case of L8 and L9 for CP, since the data points on the

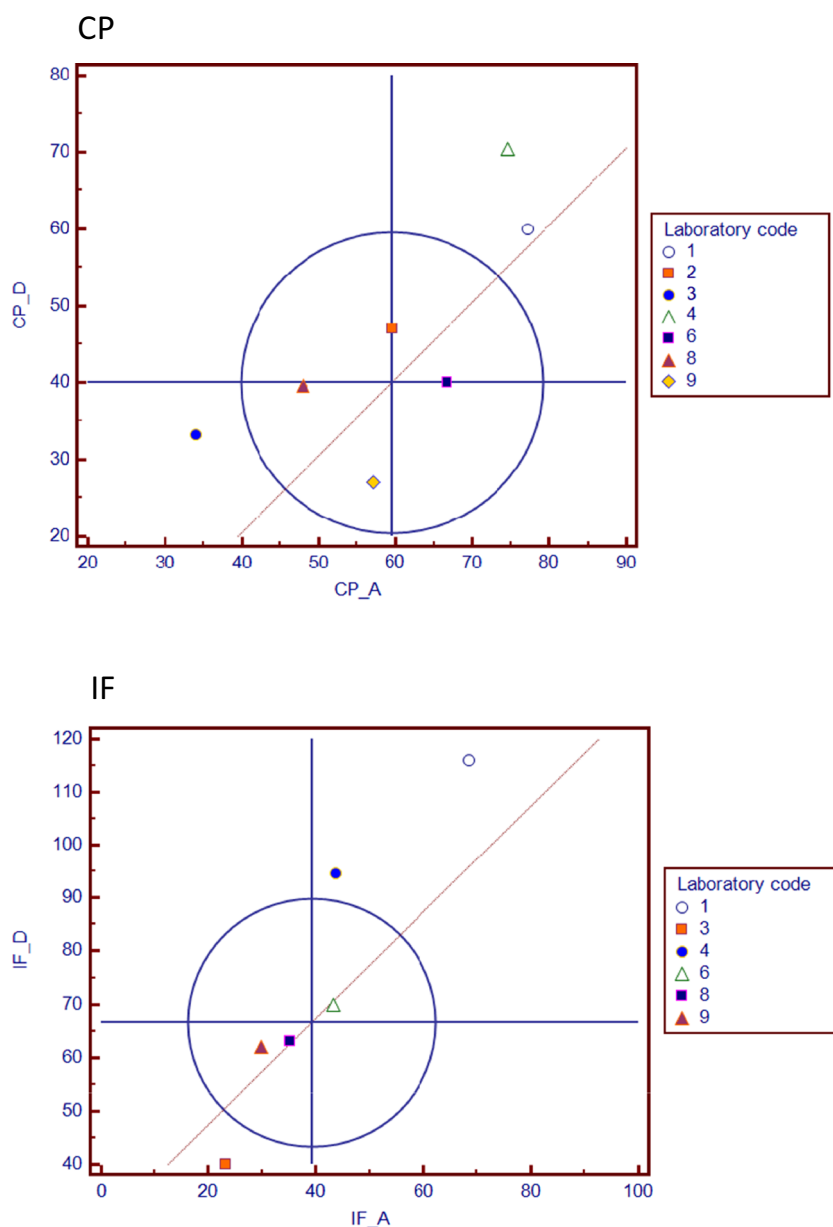
graph lie far from 45° line (Fig. 7). The points belonging to other participants (L2 and L6 for CP and L6, L8 and L9 for IF), which lie within the 95 % confidence circle and near the Manhattan median, can be labelled as acceptable (Fig. 7). Potential sources of error include the use of large calibration curve ranges (expanding from ng to $\mu\text{g L}^{-1}$) for analysis of low concentration levels, the existence of matrix interferences and matrix effects, such as signal enhancement or suppression, not corrected for with the internal standards used.

Conclusions

In conclusion, cytostatics are the least researched pharmaceuticals, which is exemplified by the low number of laboratories

participating in this study. This is expected to change as awareness grows about their presence in the environment, especially given their cytotoxic, genotoxic, mutagenic and teratogenic properties. Overall, the preparation of the samples for this inter-laboratory study was satisfactory and, with the exception of the one set of samples that was destroyed at customs, all the laboratories received their frozen samples within days. Despite low participation, a minimum amount of data were obtained for statistical analysis for the compounds CP, IF, ETO and MTX. The smallest differences between spiked values and measured values for all compounds were observed for surface water. The highest within-laboratory repeatability was observed in the analysis of the spiked hospital wastewater on account of the higher concentration of the compounds in the sample. Overall, the

Fig. 7 Youden plots for CP and IF in samples A (spiked river water) and D (spiked WWTP effluent)



reproducibility of results amongst the laboratories was poor for all compounds and matrices, with the exception of MTX in the spiked hospital wastewater. For CP and IF, the Youden test revealed in the case of surface and WWTP effluent samples three laboratories with total errors for both compounds and two laboratories with random errors for CP.

This inter-laboratory study was performed in 2013, when there were only a few laboratories performing trace analysis of cytostatics. Since then, the interest in analysing these compounds has increased significantly and we intend to repeat our inter-laboratory study within the next 2 years with a sufficient number of participant laboratories to obtain a clearer picture of the quality of analytical data obtained for a higher number of diverse cytostatic drugs. For better laboratory performance, application of deuterated internal standards for all tested compounds and MS/MS analysis will be encouraged.

Acknowledgement This work was financially supported by the EU through the EU FP7 projects CytoThreat (Fate and effects of cytostatic pharmaceuticals in the environment and the identification of biomarkers for an improved risk assessment on environmental exposure project; grant agreement No.: 265264) and SOLUTIONS (Solutions for present and future emerging pollutants in land and water resources management; grant agreement No. 603437), by the Slovenian Research Agency (Program groups P1-0143, Project L1-5457, Project J1-6744 and a Young researcher grant to M. Č.), and by the Generalitat de Catalunya (Consolidated Research Groups “2014SGR418-Water and Soil Quality Unit” and “2014SGR291-ICRA”). The work reflects only the authors’ views. The Community is not liable for any use that may be made of the information contained therein. The authors would also like to thank Prof. Walter Giger and Dr. Valeria Dulio for their support with networking and Dr. Francesco Buseti for his willingness to participate in this exercise and Dr. David Heath for providing language help.

References

- Česen M, Kosjek T, Laimou-Geraniou M, Kompare B, Širok B, Lambropoulou D, Heath E (2015) Occurrence of cyclophosphamide and ifosfamide in aqueous environment and their removal by biological and abiotic wastewater treatment processes. *Sci Total Environ* 527–528:465–473
- Crauste-Manciet S, Sessink Pj Fau - Ferrari S, Ferrari S Fau - Jomier JY, Jomier Jy Fau - Brossard D, Brossard D (2005) Environmental contamination with cytotoxic drugs in healthcare using positive air pressure isolators
- Dehouck P, Vander Heyden Y, Smeyers-Verbeke J, Massart DL, Marini RD, Chiap P et al (2003) Interlaboratory study of a liquid chromatography method for erythromycin: determination of uncertainty. *J Chromatogr A* 1010:63–74
- European Food Safety Authority. Management of left-censored data in dietary exposure assessment of chemical substances. *EFSA Journal* 2010; 8: Article 1557
- Farré M, Petrovic M, Gros M, Kosjek T, Martinez E, Heath E et al (2008) First interlaboratory exercise on non-steroidal anti-inflammatory drugs analysis in environmental samples. *Talanta* 76:580–590
- Ferrando-Climent L, Rodríguez-Mozaz S, Barceló D (2013) Development of a UPLC-MS/MS method for the determination of ten anticancer drugs in hospital and urban wastewaters, and its application for the screening of human metabolites assisted by information-dependent acquisition tool (IDA) in sewage samples. *Anal Bioanal Chem* 405:5937–5952
- Ferrando-Climent L, Rodríguez-Mozaz S, Barceló D (2014) Incidence of anticancer drugs in an aquatic urban system: from hospital effluents through urban wastewater to natural environment
- Heath E, Kosjek T, Andersen HR, Holten Lützhøft HC, Adolfsen Erici M, Coquery M et al (2010a) Inter-laboratory exercise on steroid estrogens in aqueous samples. *Environ Pollut* 158:658–662
- Heath E, Kosjek T, Farre M, Quintana JB, de Alencastro LF, Castiglioni S et al (2010b) Second interlaboratory exercise on non-steroidal anti-inflammatory drug analysis in environmental aqueous samples. *Talanta* 81:1189–1196
- Hund E, Massart DL, Smeyers-Verbeke J (2000) Inter-laboratory studies in analytical chemistry. *Anal Chim Acta* 423:145–165
- Jones TL, Betowski LD, Lopez-Avila V (1994) Validation of liquid chromatography—mass spectrometry methods: interlaboratory studies. *TrAC Trends Anal Chem* 13:333–338
- Kosjek T, Heath E (2011) Occurrence, fate and determination of cytostatic pharmaceuticals in the environment. *TrAC Trends Anal Chem* 30:1065–1087
- Kosjek T, Perko S, Žigon D, Heath E (2013) Fluorouracil in the environment: analysis, occurrence, degradation and transformation. *J Chromatogr A* 1290:62–72
- Kosjek T, Negreira N, de Alda ML, Barceló D (2015) Aerobic activated sludge transformation of methotrexate: identification of biotransformation products. *Chemosphere* 119(Supplement):S42–S50
- Llewellyn N, Lloyd P, Jürgens MD, Johnson AC (2011) Determination of cyclophosphamide and ifosfamide in sewage effluent by stable isotope-dilution liquid chromatography–tandem mass spectrometry. *J Chromatogr A* 1218:8519–8528
- Negreira N, López de Alda M, Barceló D (2013a) On-line solid phase extraction–liquid chromatography–tandem mass spectrometry for the determination of 17 cytostatics and metabolites in waste, surface and ground water samples. *J Chromatogr A* 1280:64–74
- Negreira N, Mastroianni N, López de Alda M, Barceló D (2013b) Multianalyte determination of 24 cytostatics and metabolites by liquid chromatography–electrospray–tandem mass spectrometry and study of their stability and optimum storage conditions in aqueous solution. *Talanta* 116:290–299
- Negreira N, de Alda ML, Barceló D (2014a) Cytostatic drugs and metabolites in municipal and hospital wastewaters in Spain: filtration, occurrence, and environmental risk. *Sci Total Environ* 497–498:68–77
- Negreira N, López de Alda M, Barceló D (2014b) Study of the stability of 26 cytostatic drugs and metabolites in wastewater under different conditions. *Sci Total Environ* 482–483:389–398
- Odraska P, Dolezalova L, Kuta J, Oravec M, Piler P, Blaha L (2013) Evaluation of the efficacy of additional measures introduced for the protection of healthcare personnel handling antineoplastic drugs. *Ann Occup Hyg* 57:240–250
- Parrella A, Lavorgna M, Criscuolo E, Russo C, Fiumano V, Isidori M (2014) Acute and chronic toxicity of six anticancer drugs on rotifers and crustaceans
- Van den Bossche L, Lodi A, Schaar J, Shaakov S, Zorzan M, Tranquillini ME et al (2010) An interlaboratory study on the suitability of a gradient LC-UV method as a compendial method for the determination of erythromycin and its related substances. *J Pharm Biomed Anal* 53:109–112
- Vander Heyden Y, Smeyers-Verbeke J (2007) Set-up and evaluation of interlaboratory studies. *J Chromatogr A* 1158:158–167
- Vander Heyden Y, Saevens J, Roets E, Hoogmartens J, Decolin D, Quaglia MG et al (1999) Interlaboratory studies on two high-performance liquid chromatographic assays for tylosin (tartrate). *J Chromatogr A* 830:3–28
- Yin J, Shao B, Zhang J, Li K (2010) A preliminary study on the occurrence of cytostatic drugs in hospital effluents in Beijing, China. *Bull Environ Contam Toxicol* 84:39–45

Youden WJ (1959) Graphical diagnosis of interlaboratory test results. 15, *Industrial Quality Control*, p 24–28

Web References

- Chemoth 2013; <http://chemoth.com/economics>, (Accessed 1 Aug 2013)
- CytoThreat Project 2014; <http://www.cytothreat.eu/> (Accessed 10 Aug 2013)
- Cytotoxic Safety 2013; <http://www.cytotoxicsafety.org/protect-the-environment/> (Accessed 1 Aug 2013)
- Cytotoxic Safety 2013; <http://cytotoxicsafety.org/wp-content/uploads/2014/03/Cytotoxic-Drug-Contamination-in-Surface-Water-Theresa-OKeefe.pdf> (Accessed 5 Aug 2013)
- ISO TC 69/SC 6 2011; http://www.iso.org/iso/iso_technical_committee.html%3Fcommid=49808 (Accessed 20 Jun 2013)
- ISO 13528, 2009; http://www.iso.org/iso/catalogue_detail.htm?csnumber=35664 (Accessed 2 Aug 2013)
- MedCalc Software 2015; <http://www.medcalc.org/manual/youdenplot.php> (Accessed 10 Jul 2013)
- Reuters 2008; <http://www.reuters.com/article/2008/02/19/idUS97258+19-Feb-2008+BW20080219>, (Accessed 25 Nov 2014)
- World Cancer Report 2014; B.W. Stewart, C.P. Wild (Eds.), IARC, WHO (2014). <http://edition.cnn.com/2014/02/04/health/who-world-cancer-report/> (Accessed 12 Oct 2024)