Talanta 81 (2010) 1189–1196

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Second interlaboratory exercise on non-steroidal anti-inflammatory drug analysis in environmental aqueous samples

E. Heath[a,*], T. Kosjek[a], M. Farre[b], J.B. Quintana[c,d], L.F. de Alencastro[e], S. Castiglioni[f], O. Gans[g], K. Langford[h], R. Loos[i], J. Radjenović[b], L. Mainero Rocca[j], H. Budzinski[k], D. Tsipi[l], M. Petrovic[b,m], D. Barcelo[b]

[a] *Jožef Stefan Institute, Ljubljana, Slovenia*
[b] *Department of Environmental Chemistry IDAEA-CSIC, C/ Jordi Girona 18-26, 08034 Barcelona, Spain*
[c] *IUMA - University Institute of Environment, University of A Coruña, A Coruña, Spain*
[d] *IIAA - Institute of Food Analysis and Research, Department of Analytical Chemistry, Nutrition and Food Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain*
[e] *Central Environmental Laboratory (GR-CEL), Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015-Lausanne, Switzerland*
[f] *Department of Environmental Health Sciences "Mario Negri" Institute for Pharmacological Research, Via La Masa 19, 20156 Milan, Italy*
[g] *Umweltbundesamt GmbH, Abt. Organische Analysen, Wien, Austria*
[h] *NIVA, Norwegian Institute for Water Research, 0349 Oslo, Norway*
[i] *European Commission (EC), Joint Research Centre (JRC), Institute for Environment and Sustainability, Ispra, Italy*
[j] *Universita "La Sapienza" di Roma, Italy*
[k] *Institut des Sciences Molecularies (ISM) UMR 5255 CNRS, Groupe de Physico- et Toxico-Chimie de l'environnement (LPTC), Universite Bordeaux 1, Talence, France*
[l] *Pesticide Residues Laboratory, General Chemical State Laboratory, Athens, Greece*
[m] *Institucio Catalana de Recerca i Estudis Avanzats (ICREA), Barcelona, Spain*

ARTICLE INFO

Article history:
Received 31 May 2009
Received in revised form 22 January 2010
Accepted 4 February 2010
Available online 12 February 2010

Keywords:
Non-steroidal anti-inflammatory drugs (NSAIDs)
Ibuprofen
Naproxen
Ketoprofen
Diclofenac
Interlaboratory

ABSTRACT

Several interlaboratory exercises were organised within the framework of European FP6 project NORMAN. Among others, non-steroidal anti-inflammatory drugs were investigated in different aqueous samples in two sequential ring studies. The aim of both studies was to evaluate the state-of-art in Europe and to determine possible sources of variation, while also attempting to diminish them. In the present paper we discuss the results of the 2nd Interlaboratory study, while the results of 1st round were presented before. The main scope of the 1st exercise organised within NORMAN project was to assess the laboratory proficiency regardless of the analytical method applied, to evaluate the stability of the target compounds during sample storage, and to define possible sources of variation during sample shipment, storage and analysis. In the 2nd round we primarily aimed to diminish these sources of variation by applying two predetermined analytical protocols based on liquid chromatography–mass spectrometry or gas chromatography–mass spectrometry. The two analytical protocols were compared in terms of their ability to determine individual analytes in matrices of different complexity, i.e. tap water, river water and wastewater. Furthermore, the 2nd exercise addressed also the filtration and compared the influence of different filter material categories on the analysis of non-steroidal anti-inflammatory drugs.

Results presented herein evaluate laboratory performance using *z*-score, bias, proximity and Youden plots. Overall, the laboratory performances were found to be satisfactory for determining NSAIDs in aqueous samples. The two analytical protocols, LC–MS and GC–MS, are assessed according to their sensitivity and measurement uncertainty, where the GC–MS proved superior for the analysis of Ibuprofen, Ketoprofen and Naproxen in matrices with higher complexity. Finally, neither the filtration itself, nor the filter materials were shown to significantly affect the determination of NSAIDs.

1. Introduction

Pharmaceutically active substances are a class of emerging contaminants that have raised concern in recent years. Even though the amounts of pharmaceuticals and their bioactive metabolites being introduced into environment are likely to be low, their continuous input leads to high long term presence in the environment and

* Corresponding author at: Jožef Stefan Institute, Department of Environmental Sciences, Jamova 39, 1000 Ljubljana, Slovenia. Tel.: +386 1 477 3584; fax: +386 1 251 9385.
E-mail address: ester.heath@ijs.si (E. Heath).

0039-9140/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.talanta.2010.02.009

may result in chronic effects on aquatic and terrestrial organisms. Among all pharmaceutical compounds, widespread and polar drugs such as acidic non-steroidal anti-inflammatory drugs (NSAIDs), deserve particular attention. This is due to their physico-chemical properties: high water solubility, acidic pKa, low sorption properties and often poor degradability, which allow them to pass through all man-made treatments and natural filtration steps and enter surface water, groundwater and drinking water [1,2]. The analytical methodologies for determining NSAIDs are still evolving and are applied at the level of individual laboratories. However, there is a need for harmonisation and validation of analytical methods for NSAIDs residue analysis. In the absence of standard reference materials, the main steps of analytical procedures should be evaluated by interlaboratory studies, a common practice in different research areas [3–12] and a corner stone of quality assurance [13]. The results of the analysis allow comparisons to be made between, and information to be obtained about, the laboratories, methods, or the test materials. The laboratories may come from within one organisation or may encompass different laboratories across the world. The quality value of the measurand may be known, or the object of the study may be to arrive at a consensus value. Common to all interlaboratory trials is one organisation that takes responsibility for sourcing the material, distributing it to the participating laboratories, collecting and processing of the data, and finally publishing the report [13].

Within the framework of EU FP6 project NORMAN two sequential interlaboratory studies were set-up to determine Ibuprofen, Ketoprofen, Naproxen and Diclofenac residues in various aqueous samples. The first level of verification and the quality of existing analytical procedures was evaluated by the Department of Environmental Chemistry, IDAEA-CSIC, Barcelona, Spain [8]. Here aquatic samples of differing complexity were prepared, distributed and analysed according to individual laboratory practices. The study was performed using either LC–MS/MS or GC–MS methods developed by the individual participant laboratories. All methods employed SPE as a purification and concentration step. The objectives of this 1st Interlaboratory exercise were to determine NSAIDs in different samples in various European laboratories, to evaluate the main sources of variation in the results, to ascertain any significant differences between results obtained by LC and GC for samples with different matrices, and to check the stability of the samples. The results of that study revealed the following:

(1) No significant difference between GC or LC based methods.
(2) No relation exists between the submitted results and the temperature at reception, but temperature was still recognised as an important source of variation.
(3) The number of outliers was linked to the number of steps in analytical procedure and the complexity of the method.
(4) A good agreement was obtained between the concentrations of fortification and the mean values reported by the participants.
(5) The precision of individual participants was low along the exercise suggesting the need for a protocol to unify sample treatment including handling, how to defrost the samples and chemical analysis in order to minimise sources of variation in the 2nd Interlaboratory exercise.

Based on these findings, protocols for sample pre-treatment and two harmonised protocols for GC and LC separation were prescribed. The second ring exercise was performed by the "Jožef Stefan" Institute, Ljubljana, Slovenia in collaboration with IDAEA-CSIC in Barcelona, Spain. The objectives of this study were to address weaknesses arising from the 1st Interlaboratory exercise including the influence of temperature during sample shipment, effect of sample filtration, influence of complexity of the matrix, storage of frozen cartridges instead of samples and to evaluate overall variation in results arising from LC or GC pre-defined analytical procedures.

2. Materials and methods

2.1. Experimental design, sample collection and handling

Samples were prepared by the Department of Environmental Sciences, "Jožef Stefan" Institute in collaboration with IDAEA-CSIC. Samples of wastewater treatment plant effluent (WWTP Rubi, Barcelona, Spain) and river water (Llobregat, Spain) were filtered through 2.7 μm and 0.5 μm glass micro-fibre filters. Deionised water was not filtered. Afterwards, samples were homogenised, spiked (Table 1) and sub-sampled for homogeneity and stability testing. The samples (900 mL) were transferred into 1 L polyethylene bottles and frozen. The frozen samples were then shipped on dry ice to the participating laboratories. A total number of 117 samples were sent to 13 participants in 12 laboratories, distributed among 9 European countries: Austria, France, Greece, Italy, Norway, Slovakia, Slovenia, Spain and Switzerland. The samples reached the participant laboratories within 24–72 h in a frozen state. Laboratory codes 1–13 were used to ensure anonymity. Separately, 1.5 mL of a standard NSAID mixture in methanol was sent separately at ambient temperature.

Three batches of samples were prepared for each laboratory, where each batch consisted of 3 samples prepared from wastewater (batch A), river water (batch B) and tap water (batch C) (Table 1).

2.2. Chemicals

Ibuprofen, Naproxen, Ketoprofen, and Diclofenac were supplied by Jescuder (Rubí, Spain). The purity of the standards was confirmed by LC–MS (UPLC-QTOF, Waters Corp., Milford, MA, USA) and GC–MS (Hewlett Packard, Waldbron, Germany), and the chromatographic response matched that of authentic standards purchased from Sigma Aldrich (≥98% purity, St. Louis, MO, USA). The internal standard was deuterated Ibuprofen-d3 obtained by participants themselves. *N*-Methyl-*N*-[tert-butyldimethyl-silyl]trifluoroacetamide (MTBSTFA, provided by participants) was used as a derivatising agent in the GC–MS analytical procedure.

2.3. Analytical methods

As already explained, the 1st Interlaboratory exercise [8] did not make any special requirements on the analytical procedures for determining NSAIDs. In contrast, for this round of the Interlaboratory study the participants were asked to follow the analytical protocols based on gas chromatography–mass spectrometry, GC–MS, or liquid chromatography–mass spectrometry, LC–MS, available in the participant laboratories. Both analytical procedures involved concentration and clean-up steps using an off-line solid phase extraction (SPE) with polymeric Oasis HLB 60 mg/3 mL (Waters) cartridges. Solid phase extraction was followed by LC–MS/MS or GC–MS analysis, where the latter involved an additional derivatisation step using MTBSTFA. Prior to the SPE, participants were asked to perform an additional filtration step on

Table 1
Sample matrices and encoding.

Sample code and matrix		
Batch A	Batch B	Batch C
A1 Natural wastewater	B1 Natural river water	C1 Spiked tap water
A2 Spiked wastewater	B2 Spiked river water	C2 Spiked tap water
A3 Spiked wastewater	B3 Spiked river water	C3 Spiked tap water

the sample series 1 and 2 (Table 1): A1, A2, B1, B2, C1 and C2, where the internal standard was added post-filtration. The extraction volumes were 400 mL for river water and tap water samples (batches B and C) and 200 mL for wastewater (batch A). No adjustment of the sample pH was made prior to the extraction.

2.3.1. LC–MS/MS analytical protocol

The conditioning of the SPE cartridges was carried out using 5 mL of methanol, followed by 5 mL of ultra-pure water (HPLC grade). The samples were allowed to percolate through the cartridges at a flow rate of 5 mL min^{-1}. After enrichment the cartridges were rinsed with 5 mL of HPLC grade water and then dried under vacuum (15–20 min) to remove excess water. Finally, the cartridges were eluted with 8 mL of methanol (2 × 4 mL), evaporated under nitrogen and reconstituted with 1 mL of methanol : water (25:75, v/v). The LC–MS/MS analyses were performed in negative ion mode using an RP-18 column for the chromatographic separation. The mobile phases were methanol with 5 mM NH_4 acetate and water with 5 mM NH_4 acetate. In the tandem MS operation, two multiple-reaction monitoring (MRM) transitions (identification and quantification ion) were acquired for each compound, whenever possible.

2.3.2. GC–MS analytical protocol

In the GC–MS analytical protocol the cartridges were conditioned with ethylacetate (3 mL), methanol (3 mL) and finally rinsed with ultra-pure water (3 mL). Likewise the LC–MS/MS, procedure followed the enrichment, rinsing and drying step, after which the cartridges were eluted with 2 × 1 mL of ethylacetate. Prior to GC–MS analysis the samples were derivatised with MTBSTFA at 60 °C for 1 h. Separation was performed using (30 m × 0.25 mm × 0.25 μm) capillary column with 95% methyl/5% phenyl polysiloxane stationary phase for chromatographic separation. The GC oven was programmed as follows: 2 min at 65 °C, then ramped at 30 °C/min to 180 °C, further ramped at 5 °C/min to 300 °C, and finally held for 12 min. The target ions used for quantification were m/z 263 for Ibuprofen, m/z 287 for Naproxen, m/z 311 for Ketoprofen and m/z 352 and 354 for Diclofenac.

2.4. Statistical parameters

The homogeneity of sample preparation was statistically evaluated by a χ^2 test, using Eq. (1):

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

where O_i is a mean concentration of two parallels in each sample and E_i the mean concentration of each batch containing 10 samples.

As an acceptance criterion for each result the z-score value was calculated using Eq. (2):

$$z = \frac{x_{lab} - x_0}{\sigma_0} \quad (2)$$

where x_{lab} is a laboratory mean (median), x_0 a mean (median) concentration for the initial dataset and σ_0 is the corresponding standard deviation. Both, the classical approach using mean value and the robust approach using the median were employed in calculations of z-scores. A z-score higher than 3.0 indicates an unsatisfactory performance [14] and such result was automatically excluded from the further data processing as an outlier. For values between $2.0 \le |z| \le 3.0$ [7], that is suspect outliers, the Dixon discordance test [15] was applied at a 5% significance level.

After excluding any outliers, the corrected mean (X), standard deviation (σ), coefficient of variation (CV), standard error of mean (σ_M), median (M), minimum (Min) and maximum value (Max) were calculated for each series of results and the normality was confirmed by Lilliefor's Test. In addition, laboratory biases (D) were estimated for each result (or average of results) reported by each participant. According to ISO/DIS 13528:2002(E) [16] the biases were classified into three categories $|D \ge 3.0\sigma|$ indicating an "action signal", $|2.0\sigma \le D < 3.0\sigma|$ considered as a "warning signal" and $-2.0\sigma < D < 2.0\sigma$ indicating "acceptable value". The outlier results were excluded from the calculation of D.

Proximity to the mean Pr(X) was calculated as:

$$\Pr(X) = \frac{1}{n} \times \sum \frac{|x_{lab} - X|}{X} \quad (3)$$

where n is the sample size, x_{lab} the laboratory mean and X the corrected mean after the outlier exclusion. Similarly, the 'proximity to the median' Pr(M) was calculated, as shown in Eq. (4).

$$\Pr(M) = \frac{1}{n} \times \sum \frac{|x_{lab} - M|}{M} \quad (4)$$

To evaluate the effect of filtration on determination of NSAIDs in different matrices three statistical tests were used. An F-test at 5% significance level was used for comparison of the variances between filtered (series '2') and unfiltered (series '3') sample series within each batch [17]. In addition, a paired t-test was applied to compare unfiltered and filtered mean values within each laboratory. In case of Ibuprofen in tap water samples (batch C, series, 1, 2 and 3) three variances were compared with ANOVA.

3. Results and discussion

3.1. Sample preparation

To minimise the variation between the samples, the matrices were collected, prefiltered, spiked, homogenised, divided-up and frozen within 24 h. To assure and to confirm the quality of sample preparation, the homogeneity of mixing was tested for each batch (matrix). Thus the homogenised batches (A, B and C, series 2 and 3, Table 1), were spiked and then sub-sampled. According to ISO/DIS 13528 [16] 10 samples per batch (series 2 and 3) were taken randomly from different layers in the polyethylene container. The samples were then analysed in parallels and the homogeneity was statistically evaluated using the χ^2 test. The homogeneity was confirmed in all cases at the 95% confidence level.

Given that the stability of samples during shipment and storage was one of the goals of the 1st NORMAN Interlaboratory exercise [8], and was confirmed for all analytes, the authors felt that there was no need to perform similar tests in the 2nd round. Furthermore, testing the stability of compounds in the aqueous media was an irrelevant issue for 2nd round as participants were asked to perform the SPE extraction short upon sample receipt (48 h), while the analyses themselves were allowed to be carried out within 3 months from extraction. Instead, the stability of NSAIDs in frozen cartridges was tested within 3 months after the sample extraction, where no decrease in the analyte content was observed within the studied period of time.

3.2. Laboratory proficiency testing

A total number of 108 samples were analysed in this exercise by 12 participants from 11 different institutions. Seven LC and five GC laboratories took part in the exercise and submitted 773 results including parallel and <LOD values. Among these, 428 values were pooled out for subsequent data mining process, starting with the determination of outliers. The z-score calculation, which was performed by classical and robust approach and the Dixon test, yielded 15 (3.5%) or 18 (4.2%) outliers, respectively. Fig. 1 shows the absolute z-score values according to the classical approach for each laboratory.

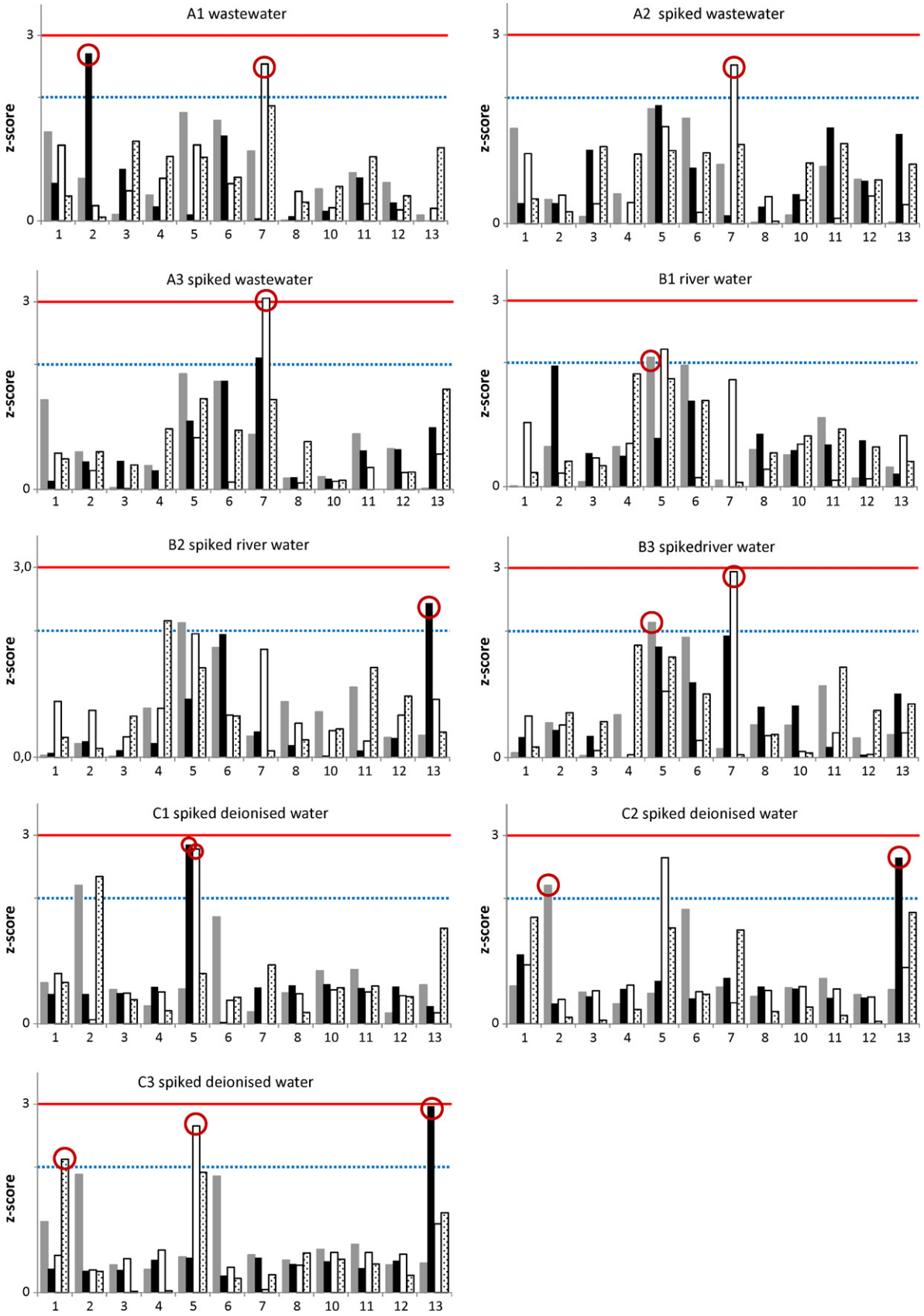

Fig. 1. Absolute *z*-score values for each of the participant laboratories (Lab ID 1–13) calculated using classical approach. The outliers are marked with circles. Dotted line: $z = 2$; solid line: $z = 3$.

The highest number of outliers was determined for tap water, which is a consequence of the lower concentration levels, agreeing with the results of the 1st Interlaboratory exercise [8]. By the use of the robust approach, the total number of outliers increased for three, which were, as shown in Table 2, all on account of Ibuprofen in tap water using the LC–MS/MS analytical procedure. Out of 249 results obtained by LC–MS/MS, 12 (4.8%) or 15 (6.0%) were excluded as outliers using the classical and robust approach, respectively. The number of excluded GC results was considerably lower compared to the LC method, i.e. 3 out of 179 (1.7%), regardless of which outlier-testing approach was used. Out of the total of twelve participants, the outlier results were reported by 5 (or 6 using the robust approach) laboratories, which generally experienced difficulties in determining only a single compound (e.g. Naproxen in laboratory 7 and Ketoprofen in laboratory 13), while, on the other hand, they showed a satisfactorily method performance for determining the remaining analytes (Fig. 1).

The percentage of the outlier values reported in the 2nd round was considerably lower compared to the 1st round [8], which may be attributed to diminishing the weaknesses recognised in the 1st Interlaboratory exercise, i.e. sample shipment to the participant laboratories and/or pre-determination of the analytical protocols. Unifying the analytical protocols revealed a particularly evident improvement in case of GC based analytical protocol, where the number of outliers was up to 5 folds lower in 2nd round (3 outliers out of 428 processed results), when compared to 1st round (15 outliers out of 486 results). Such outcome is in agreement with our expectations since GC analytical protocols employed in the 1st exercise were rather heterogeneous, i.e. up to 3 different sorbent materials for SPE, 3 different elution solvents and 4 different derivatisation reagents were used [8].

After the outlier exclusion, the mean, standard deviation, coefficient of variation, standard error of mean, median, minimum and maximum value were calculated for each NSAID in each of the 9 samples. As shown in Table 3, the sample series '2' and '3' were spiked with the same level of the tested compounds, while spiking levels of Ibuprofen in tap water were equal in all three series: '1', '2' and '3'. An excellent agreement was obtained between the concentrations measured in tap water and the actual spiking levels (Table 3). In general, the mean and median concentrations matched closely and followed the spiked concentrations. However, in wastewater and river water, the concentration raise is, as a rule, smaller than the actual concentration increase due to spiking. Clearly, this phenomenon can issue from the suppression effect of the matrix compounds, which interfere with the LC and GC analysis. This effect is principally emphasized in the atmospheric pressure ionisation methods (API), which are generally used in LC, and have been several times reported [18–23]. Thus, it was shown that, in batch B the LC samples yielded a considerably lower response to the spiked concentration in comparison to the GC samples, which were in general only slightly affected. This effect corresponds to complexity of applied river water sample (batch B), which in our case contained a lot of particulate matter and was sampled during the dry season downstream the city of Barcelona. In addition, it was extracted in twice higher volume (400 mL) as wastewater (200 mL) and thus the extracts presumably involved more organic matter than wastewater. Further, an unusually high concentration of Ibuprofen (approx. 7.5 $\mu g\,L^{-1}$, Table 3) was determined in the natural river water sample, what is more likely a result of the sampling conditions (grab sample, extremely low water level and flow at sampling time), than a picture of Llobregat River pollution.

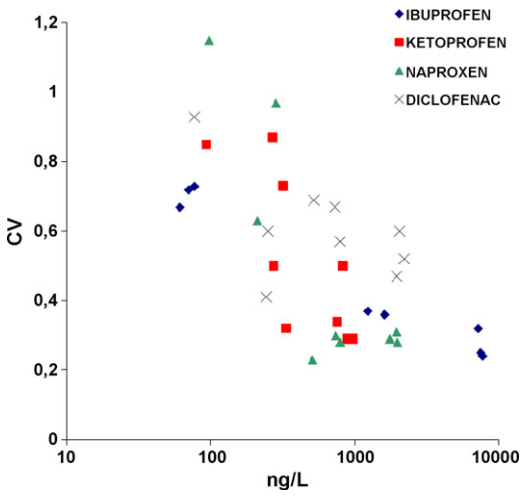

Fig. 2. CV plotted against the mean concentrations of tested compounds determined in wastewater, river water and tap water samples.

Alternatively, high coefficients of variation (CV, Table 3) were observed particularly in tap water for all four analytes due to the low spiking levels, while Diclofenac revealed high variability also in the remaining two matrices. As further discussed, this indicates that the highest uncertainty was found in the determination of Diclofenac, without respect to the analytical method used. Additionally, the smallest number of outliers observed for this compound is also attributed to the high coefficient of variation. Indeed, as shown in Fig. 2, the CV increases for the lowest concentration samples (tap water), but remains lower for Ibuprofen, probably as a consequence of the fact that the internal standard is Ibuprofen-d3. Therefore, a more suitable internal standard (e.g. isotopically labelled Diclofenac) may improve method performance for Diclofenac.

A comparison of the CV values (Fig. 3) in both complex matrices, i.e. river water and wastewater, resulted in considerably higher CV in the samples analysed by LC–MS procedure, which again implies that the ion suppression affected the LC analysis. On the other hand, the GC–MS provided a smaller variability in the analysis of Ibuprofen, Ketoprofen and Naproxen in both matrices, whereas, the determination of Diclofenac did not prove particularly consistent, regardless of the analytical procedure.

The ISO/DIS 13528 [16] classification of the laboratory biases resulted in the complete absence of "action signals" and one or less "warning signals" per series of results, which indicated that the calculated mean (X) and standard deviation (σ), with the underlying normal distribution, were good approximates for the true mean and the standard deviation values. The Pr(X) and Pr(M) values were derived for each analyte determined by the participating laboratories. These parameters describe a general capability of a laboratory to determine an analyte without respect to the tested matrix. Pr(X)

Table 2
Calculated number (the percentage of total is stated in the brackets) of outlier values according to the sample matrix, analyte and analytical method. The outliers were calculated according to the classical and robust approach.

	Classical approach		Robust approach
Outliers/sample matrix			
Wastewater	4 (0.9%)		4 (0.9%)
River water	4 (0.9%)		4 (0.9%)
Tap water	7 (1.6%)	+3	10 (2.3%)
Outliers/analyte			
Ibuprofen	3 (0.7%)	+3	6 (1.4%)
Ketoprofen	5 (1.2%)		5 (1.2%)
Naproxen	6 (1.4%)		6 (1.4%)
Diclofenac	1 (0.2%)		1 (0.2%)
Outliers/analytical method			
LC–MS/MS	12 (2.8%)	+3	15 (3.5%)
GC–MS	3 (0.7%)		3 (0.7%)

Table 3
Summary of the corrected statistical parameters after the outlier exclusion: mean (X), standard deviation (σ), coefficient of variation (CV), standard error of mean (σ_M), median (M), minimum (Min) and maximum (Max).

Sample	Matrix	Spiking level ($ng\,L^{-1}$)	Filtration	No. acc. results	X ($ng\,L^{-1}$)	σ ($ng\,L^{-1}$)	CV	σ_M ($ng\,L^{-1}$)	M ($ng\,L^{-1}$)	Min ($ng\,L^{-1}$)	Max ($ng\,L^{-1}$)	No. outliers
Ibuprofen												
A1	Wastewater	–	YES	12	1238	460	0.37	133	1265	433	1987	0
A2		416	YES	12	1622	577	0.36	167	1668	570	2588	0
A3		416	NO	12	1620	586	0.36	169	1669	537	2633	0
B1	River water	–	YES	11	7545	1853	0.25	559	7351	4500	11,684	1
B2		416	YES	12	7250	2302	0.32	665	7537	2358	11,235	0
B3		416	NO	11	7791	1864	0.24	332	7663	4600	11,891	1
C1	Tap water	50	YES	12	77	56	0.73	16	55	29	200	0
C2		50	YES	11	61	41	0.67	12	46	33	172	1
C3		50	NO	12	70	50	0.72	14	47	31	571	0
Ketoprofen												
A1	Wastewater	–	YES	11	334	108	0.32	33	350	111	520	1
A2		790	YES	12	967	284	0.29	82	985	434	1400	0
A3		790	NO	12	830	416	0.50	120	905	107	1705	0
B1	River water	–	YES	10	269	234	0.87	74	147	69	725	0
B2		790	YES	11	754	259	0.34	78	812	91	997	1
B3		790	NO	12	886	261	0.29	75	893	428	1389	0
C1	Tap water	47	YES	11	93	79	0.85	24	40	30	217	1
C2		205	YES	11	319	231	0.73	70	248	123	854	1
C3		205	NO	11	273	136	0.50	41	230	170	571	1
Naproxen												
A1	Wastewater	–	YES	11	507	115	0.23	35	510	325	675	1
A2		412	YES	11	791	224	0.28	67	808	332	1022	1
A3		412	NO	11	737	220	0.30	66	742	317	1030	1
B1	River water	–	YES	12	1754	516	0.29	149	1825	609	2646	0
B2		412	YES	12	1956	608	0.31	175	1976	771	2993	0
B3		412	NO	11	1978	563	0.28	170	1977	852	2925	1
C1	Tap water	45	YES	10	97	111	1.15	35	46	26	388	1
C2		120	YES	12	283	276	0.97	80	154	113	1014	0
C3		120	NO	11	210	132	0.63	40	167	111	516	1
Diclofenac												
A1	Wastewater	–	YES	12	521	357	0.69	103	586	59	1186	0
A2		523	YES	12	730	487	0.67	141	693	110	1341	0
A3		523	NO	11	796	452	0.57	136	860	71	1444	0
B1	River water	–	YES	12	1959	924	0.47	267	1887	352	3640	0
B2		523	YES	12	2054	1234	0.60	356	2030	300	4715	0
B3		523	NO	12	2216	1152	0.52	332	2284	386	4262	0
C1	Tap water	63	YES	12	77	71	0.93	21	48	10	243	0
C2		220	YES	12	250	149	0.60	43	245	22	515	0
C3		220	NO	11	244	101	0.41	30	233	21	433	1

and Pr(M) are plotted in Fig. 4a and b, respectively, where the outlier values are included in order to illustrate a full performance of laboratories [7]. Comparison of both figures reveals that Fig. 4b, where Pr(M) is plotted for each laboratory, more significantly shows the differences in laboratory performance regarding each test compound. In addition, on the x-axes the analytical protocol is marked, which, in contrast with the results from the first exercise, shows a relatively good performance of GC laboratories and may be the consequence of method harmonisation, as explained before.

Fig. 4 also shows that one laboratory in particular has a high Pr(X) and Pr(M) values for three compounds, while the six laboratories (3, 4, 8, 10, 11 and 12) showed excellent method performance for all four analytes.

As C1 and C2 were split-level samples with respect to the concentration of Ibuprofen (Table 3), a Youden graph was plotted from the reported results for C1 samples (x-axes) against those reported for C2 samples (y-axes, Fig. 5). The median values for both samples are also plotted (dotted lines: $x = 55\,ng\,L^{-1}$, $y = 46\,ng\,L^{-1}$), where their intersection point is accepted as the most probable value [7]. The results show an excellent agreement with the spiking level for Ibuprofen in the batch C, i.e. $50\,ng\,L^{-1}$.

Further, the three isolated points positioned in the upper right quadrant of the Youden plot (Fig. 3) illustrate that the reported results were too high indicating a systematic error in these three laboratories [7,13]. Youden plots were plotted for all tested compounds in at least two series of samples. Since all the plots reveal similar outcomes only one representative plot is shown.

3.3. Filtration test

Filtration as a step in sample preparation process may cause two additional effects. First, depending on the analyte polarity and filter material, the analytes can adsorb to a filter, and consequently the concentrations determined in final samples are lower than the actual concentrations before filtration. On the other hand, by removing the organic matter present in matrix, the filtration may be a way to reduce the ion suppression effect and thus improve the LC–MS analysis.

In order to evaluate the effect of filtration the sample series '2' and '3' in each batch (A, B, C) were prepared in parallel. As indicated in Table 3, the participants were asked to filter samples in series '2' prior to SPE, while sample in series '3' were extracted without the pre-filtration. By statistical testing of the reported values it was shown that the samples were drawn from the same population, suggesting that the filtration had no effect on the analysis. As illustrated in Table 4, only in case of Naproxen in tap water a significant difference in variances was observed. However, this was not confirmed with the 'paired t-test' on sample means and therefore it is concluded that filtration did not cause the difference in two-sample variances.

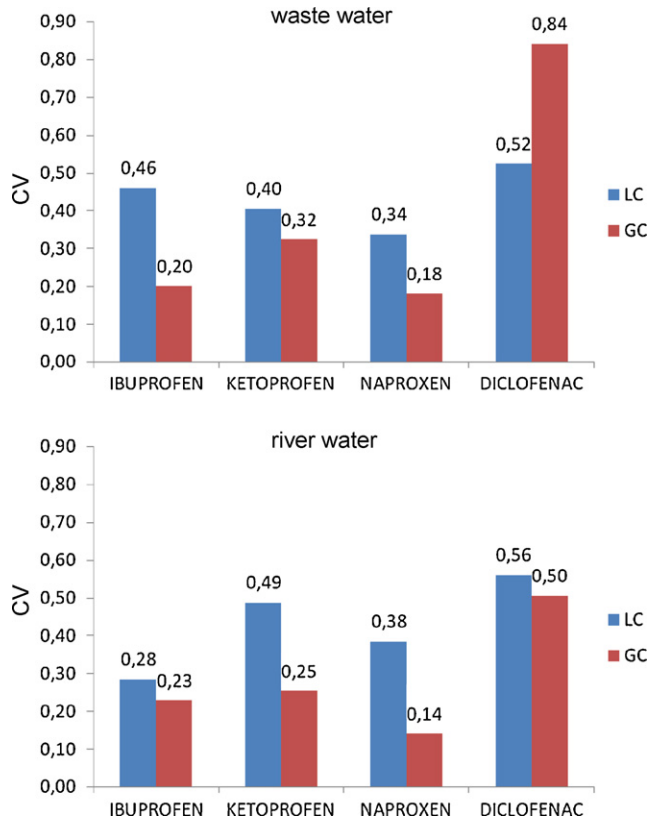

Fig. 3. CV derived for LC–MS and GC–MS procedure. Top: wastewater; bottom: river water.

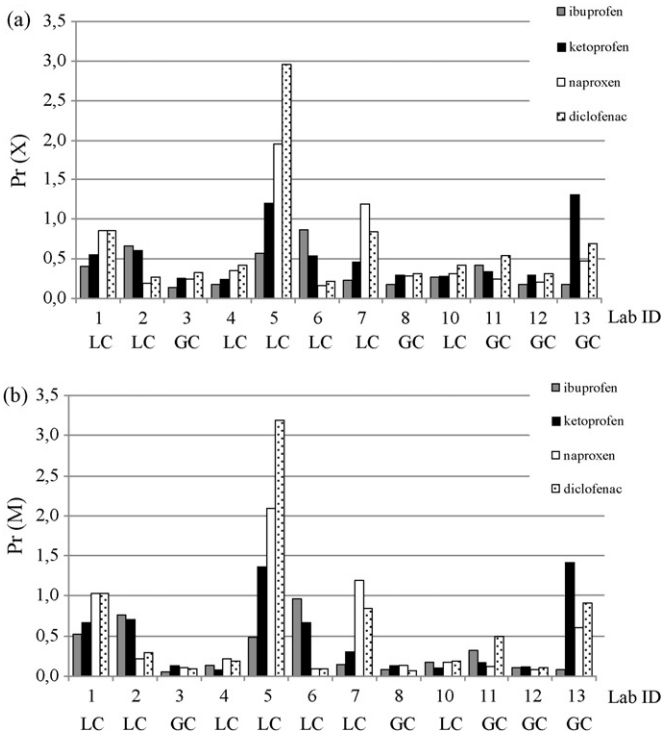

Fig. 4. (a) Bar-chart showing the Pr(X) values for each analyte measured by participating laboratories. (b) Pr(M) results. The analytical procedure is indicated bellow laboratory ID.

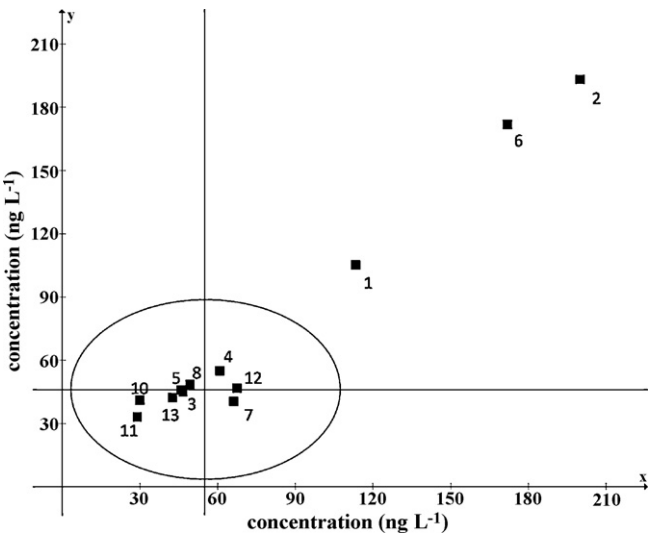

Fig. 5. Two-sample Youden plot for Ibuprofen in C1 and C2 samples. The borderline represents 95% around the origin of the plot.

Table 4
Results on testing the effect of filtration.

Tested samples	Test significance		
	F-test	*t*-test	ANOVA
Ibuprofen			
A2/A3	NO	NO	
B2/B3	NO	NO	
C1/C2/C3			NO
Ketoprofen			
A2/A3	NO	NO	
B2/B3	NO	NO	
C2/C3	NO	NO	
Naproxen			
A2/A3	NO	NO	
B2/B3	NO	NO	
C2/C3	YES	NO	
Diclofenac			
A2/A3	NO	NO	
B2/B3	NO	NO	
C2/C3	NO	NO	

As the filter material was not predetermined in the analytical protocols at least four different types of materials were used in different laboratories: glass-fibre, nitrocellulose membrane, nylon membrane, cellulose acetate and non-specified membrane filters. Among the twelve participating laboratories, seven used glass micro-fibre filters, while the remaining five used membrane filters. In order to test the influence of the filter material, *F*-test was applied to compare the variances between glass-fibre and membrane filtering in all the filtered samples (A1, A2, B1, B2, C1 and C2). It was proven that at a 95% confidence level that the filter material had no impact on the analysis of NSAIDs.

By showing that filtration does not impact determination of the compounds that is no sorption on the filters was observed and the matrix effect was not reduced, the series of samples 2 and 3 (and all three series in case of Ibuprofen in tap water) were made possible to be compared on measures of precision. Obtained results are in agreement with published literature [24].

4. Conclusions

Twelve participants from eleven different European research institutes and universities took part in a 2nd Interlaboratory exer-

cise on the determination of selected NSAIDs in aqueous matrices. The 1st NORMAN Interlaboratory exercise was a test round focusing on the stability of compounds during sample storage, whereas the 2nd round was based on two predetermined analytical protocols (LC–MS/MS and GC–MS). Further, the 2nd round specifically addressed the filtration and eliminated the weaknesses recognised in the 1st round. Thus, in contrast to the 1st round, the samples were shipped on dry ice and were extracted on arrival at the participating laboratories.

On the basis of the 1st and 2nd Interlaboratory exercise we conclude that shipping samples on dry ice, as well as using a standard laboratory protocol contributed towards a reduced number of outliers and improved laboratory performance, particularly for GC analysis. Thus, the distribution of the outliers between the GC and LC protocols is contrary to the results of the 1st round of the NORMAN Interlaboratory exercise. However, as the outliers were distributed among only 5 participants this suggests that the performance of a single laboratory had a large impact on the final number of outliers. Another aim of the 2nd round was to test, whether the pre-filtration affected the determination of the analytes in the tested matrices. The results of the test implied that the filtration itself as well as filter material, did not affect the analysis of selected NSAIDs in none of the three tested matrices.

Within the 2nd round the two analytical protocols, LC–MS and GC–MS, are assessed according to their sensitivity and measurement uncertainty. On the basis of the results which included 7 LC based and 5 GC based results, GC–MS analytical procedure was proved superior for the analysis of Ibuprofen, Ketoprofen and Naproxen in matrices with higher complexity. Higher uncertainty was found in the determination of Diclofenac, without respect to the analytical method used. To verify the outcomes of this Interlaboratory exercise an option would be to involve the higher number of participants.

Importantly, the results of the 2nd Interlaboratory exercise are not directly comparable with the 1st Interlaboratory exercise, especially not in terms of repeatability and reproducibility of results. The main reason is that in the 1st Interlaboratory exercise different matrices were spiked with the same amounts of analytes in order to confirm stability, while in the 2nd Interlaboratory round spiking and/or treatment of samples differed.

Acknowledgements

This work has been supported by EU FP6 project NORMAN (contract no.: 018486) and Ad Futura grant (523-026/2006). Authors wish to thank Mrs. Silva Perko for her valuable advices and support in sample preparation.

This article reflects only author's views and the EU is not liable for any use that may be made of the information contained herein.

References

[1] T. Kosjek, E. Heath, B. Kompare, Anal. Bioanal. Chem. 387 (4) (2007) 1379–1387.
[2] T. Kosjek, D. Žigon, B. Kralj, E. Heath, J. Chromatogr. A 1215 (1–2) (2008) 57–63.
[3] A. Marchetto, R. Mosello, M. Bianchi, H. Geiss, G. Serini, G. Serrini Lanza, G.A. Tartari, H. Muntau, Microchim. Acta 123 (1–4) (1996) 95–102.
[4] R.R. Maronpot, K. Mitsumori, P. Mann, M. Takaoka, S. Yamamoto, T. Usui, H. Okamiya, S. Nishikawa, T. Nomura, Toxicology 146 (2000) 149–159.
[5] U. Kalbe, W. Berger, F.-G. Simon, J. Eckardt, G. Christoph, J. Hazard. Mater. 148 (2007) 714–720.
[6] M. Ricci, O. Bercaru, R. Morabito, C. Brunori, I. Ipolyi, C. Pellegrino, A. Sahuquillo, F. Ulbreth, TrAC Trends Anal. Chem. 26 (2007) 818–827.
[7] Y.V. Heyden, J. Smeyers-Verbeke, J. Chromatogr. A 1158 (2007) 158–167.
[8] M. Farré, M. Petrovic, M. Gros, T. Kosjek, E. Martinez, E. Heath, P. Osvald, R. Loos, K. Le Menach, H. Budzinski, F. De Alencastro, J. Müller, T. Knepper, G. Fink, T.A. Ternes, E. Zuccato, P. Kormali, O. Gans, R. Rodil, J.B. Quintana, F. Pastori, A. Gentili, D. Barcelo, Talanta 76 (2008) 580–590.
[9] E. Heath, T. Kosjek, H.R. Andersen, H.-C. Holten Lützhøft, M. Adolfson Erici, M. Coquery, R.-A. During, O. Gans, C. Guignard, P. Karlsson, F. Manciot, Z. Moldovan, D. Patureau, I. Cruceru, F. Sacher, A. Ledin, Proceedings of "XENOWAC – Xenobiotics in Urban Water Cycle": Interlaboratory Exercise on Steroid Estrogens in Aqueous Samples, Cyprus, 11–13 March, 2009.
[10] H.W. Vesper, S. Bhasin, C. Wang, S.S. Tai, L.A. Dodge, R.J. Singh, J. Nelson, S. Ohorodnik, N.J. Clarke, W.A. Salameh, C.R. Parker jr, R. Razdan, E.A. Monsell, G.L. Myers, Steroids 74 (2009) 498–503.
[11] F.T. Peters, O.H. Drummer, F. Musshoff, Forensic Sci. Int. 165 (2007) 216–224.
[12] I. Taverniers, M. De Loose, E. Van Bockstaele, TrAC Trends Anal. Chem. 23 (2004) 535–552.
[13] D.B. Hibbert, in: P. Worsfold, A. Townshend, C. Poole (Eds.), Encyclopedia of Analytical sciences, vol. 7, Elsevier Ltd., Oxford, 2005, p. 449.
[14] M. Thompson, R. Wood, Pure Appl. Chem. 65 (9) (1993) 2123.
[15] S.P. Verma, A. Quiroz-Ruiz, Revista Mexicana de Ciencias Geológicas 23 (2) (2006) 133–161.
[16] International Organization for Standardization (2002), Statistical methods for use in proficiency testing by interlaboratory comparisons, Draft International Standard ISO/DIS 13528.
[17] M. Pečanac, V. Hudnik, Merilna negotovost pri kemijskih meritvah, ZTI, Ljubljana, March 12–13, 2008.
[18] W.W. Buchberger, Anal. Chim. Acta 593 (2007) 129–139.
[19] T.A. Ternes, TrAC Trends Anal. Chem. 20 (2001) 419–434.
[20] M. Petrović, E. Eljarrat, M.J. Lopez de Alda, D. Barceló, J. Chromatogr. A 974 (2002) 23–51.
[21] M. Petrović, M.D. Hernando, M.S. Diaz-Cruz, D. Barceló, J. Chromatogr. A 1067 (2005) 1–14.
[22] J.P. Antignac, K.D. Wasch, F. Monteau, H.D. Brabander, F. Andre, B.L. Bizec, Anal. Chim. Acta 529 (2005) 129–136.
[23] A. Kloepfer, J.B. Quintana, T. Reemtsma, J. Chromatogr. A 1067 (2005) 153–160.
[24] I. Rodriguez, J.B. Quintana, J. Carpinteiro, A.M. Carro, R.A. Lorenzo, R. Cela, J. Chromatogr. A 985 (2003) 265–274.