Environmental Research 298 (2026) 124218

Contents lists available at ScienceDirect

Environmental Research

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

Effect-based methods and effect-based trigger values for estrogenicity monitoring in surface water: an interlaboratory study

Livia Gómez [a,1], Dimitar Marinov [b,1], Elena Porcel-Rodríguez [c], Diletta Scaccabarozzi [c], Isabella Sanseverino [a,1], Pietro Sciuto [c], Frank Narendja [d], Romana Hornek-Gausterer [d], Katharina Lenz [d], Carole Chalon [e], Yves Marneffe [e], Přemysl Soldán [f], Martin Hora [f], Matti t. Leppänen [g], Johanna Järvistö [g], Jyrki Viidanoja [g], Selim Ait-Aissa [h], Emmanuelle Maillot-Marechal [h], Abd El Rahman El Mais [h], Beate I. Escher [i], Maria König [i], Henner Hollert [j], Simone Wollenweber [j], Sebastian Buchinger [k], Marina Ohlig [k], Roberta Pedrazzani [l], Johan Lundqvist [m], Agneta Oskarsson [m], Etienne Vermeirssen [n], Sibylle Maletz [n], Daniel Olbrich [n], Alvine C. Mehinto [o], Kameron Wong [o], Nancy Denslow [p], Luca Amendola [q], Ruben Gil-Solsona [r], Pablo Gago-Ferrero [r], Monica Potalivo [s], Mario Carere [t], Teresa Lettieri [c,*,2]

[a] *Piksel S.r.l., Italy*
[b] *ARHS Developments S.A., 4370, Belvaux, Luxembourg*
[c] *European Commission, Joint Research Centre (JRC), Ispra, Italy*
[d] *Environment Agency Austria (EAA), Spittelauer Laende 5, 1090, Vienna, Austria*
[e] *Scientific Institute of Public Service (ISSEP), Rue Du Chéra, 200, 4000, Liège, Belgium*
[f] *T. G. Masaryk Water Research Institute (VUV TGM)- Hydrobiology Lab, Macharova 954/5, 702 00, Ostrava, Czech Republic*
[g] *Finnish Environment Institute (SYKE), Survontie 9A, FI-40500, Jyväskylä, Finland*
[h] *Institut National de L'Environnement Industriel et des Risques (INERIS), Technologique ALATA - BP2, 60550, Verneuil-en-Halatte, France*
[i] *Department of Cell Toxicology. Helmholtz Centre for Environmental Research (UFZ), Permoserstraße 15, 04318, Leipzig, Germany*
[j] *Goethe Universität Frankfurt, Max-von-Laue-Str. 13, 60438, Frankfurt Am Main, Germany*
[k] *Federal Institute for Hydrology (BfG), Am Mainzer Tor 1, 56068, Koblenz, Germany*
[l] *Dipartimento di Ingegneria Meccanica e Industriale (DIMI), Via Branze, 38, 25123, Brescia, Italy*
[m] *Swedish University of Agricultural Sciences, Department of Animal Biosciences, Ulls Väg 26, SE-750 07, Uppsala, Sweden*
[n] *Swiss Centre for Applied Ecotoxicology, Überlandstrasse 133, CH-8600, Dübendorf, Switzerland*
[o] *Southern California Coastal Water Research Project, 3535 Harbor Blvd, Suite 110, Costa Mesa, CA, 90815, USA*
[p] *Center for Environmental and Human Toxicology, University of Florida, 2187 Mowry Road, Gainesville, FL, 32611, USA*
[q] *ARPA Lazio, The Regional Environmental Protection Agency-Lazio Region, Italy*
[r] *IDAEA-CSIC, Carrer de Jordi Girona, 18-26, 08034, Barcelona, Spain*
[s] *ISPRA - Environmental Metrology Unit, Via Del Fosso di Fiorano 64, 00143, Rome, Italy*
[t] *National Health Institute (ISS), Viale Regina Elena 299, 00161, Rome, Italy*

A R T I C L E I N F O

Keywords:
Bioassays
Safety thresholds
Water quality monitoring
Chemical mixtures
Hormones
Endocrine-disrupting chemicals

A B S T R A C T

Effect-based methods (EBMs) may be included in the European Water Framework Directive (WFD) to evaluate estrogenic substances. The European Commission's Joint Research Centre conducted an interlaboratory study to assess estrogenic EBMs and effect-based trigger (EBTs) values derived using three options: (1) linking the EBT value to environmental quality standards (EQS), (2) correlating *in vitro* and *in vivo* data, and (3) averaging bioassay-specific EBT values. Surface water samples from eight Northern-Italian sites containing estrogenic hormones and endocrine-disrupting chemicals (EDCs) were analysed by fourteen laboratories employing EBMs, while four laboratories performed chemical analysis. Chemical data indicated cumulative risk in several samples,

* Corresponding author.
E-mail address: teresa.lettieri@ec.europa.eu (T. Lettieri).
[1] The work was carried out at/on behalf of the European Commission, Joint Research Centre (JRC), Ispra, Italy.
[2] Postal address: Via E. Fermi 2749, 21027 Ispra (VA), Italy.

https://doi.org/10.1016/j.envres.2026.124218
Received 16 January 2026; Received in revised form 26 February 2026; Accepted 6 March 2026
Available online 7 March 2026

with estrone and bisphenol A identified as main drivers. All EBMs detected estrogenic activity, but specificity differed: some bioassays responded mainly to hormones, whereas others also responded to non-hormonal EDCs. EBMs flagged estrogenic risk in a sample that showed no individual exceedances of EQS. Applying EBT option 1 yielded the highest concordance with chemical results, achieving full compliance in eight bioassays and proven to be the most protective. Indeed, option 2 reduced the risk quotient (RQ) by > 30%, leading to two bioassays in full compliance, while Option 3 resulted in RQ changes (<20%) for most EBMs, with seven bioassays in full compliance with chemical analysis.

The study underscores the need to harmonise EBMs – including data evaluation – to address chemical mixtures and provides recommendations for Member States on their application in the WFD. Integrating EBMs with conventional monitoring enhances protection against cumulative estrogenic risks from both hormones and EDCs.

Abbreviations list:

AF –	Assessment Factor
BEQ -	bioanalytical-equivalent concentration
BPA -	Bisphenol A
DEHP -	di(2-ethylhexyl) phthalate
E1 -	Estrone
E2 -	17β-estradiol
EE2 -	17α-ethinylestradiol
EBM -	Effect-based methods
EBT -	Effect-based trigger value
EDCs -	Endocrine-disrupting chemicals
EDCM -	Endocrine-disrupting chemicals mixture
EEQ -	E2-equivalent concentration
EH -	Estrogenic hormones
EM -	Estrogenic hormones mixture
EQS -	Environmental Quality standards
MoA -	Mode of Action
4-NP -	4-nonylphenol
PS -	Priority Substances
RQ –	Risk quotient
SOP -	standard operating procedure
SP -	Sampling Point
WFD -	Water Framework Directive
WS -	Water Sample

1. Introduction

Thousands of chemicals enter aquatic environments from different sources, including industrial discharges, agricultural runoff, sewage treatment plants, and urban stormwater (Singh et al., 2024). These substances can interact, potentially leading to greater environmental hazards than individual chemicals alone, posing significant concerns to both environmental and human health (Luijten et al., 2023; Savitz and Hattersley, 2023). Among them, hormones and endocrine-disrupting chemicals (EDCs) are particularly concerning. EDCs can interfere with the hormonal system of humans and animals, while estrogenic hormones (EH) are especially critical due to activity at very low (pg/L or ng/L) concentrations (Könemann et al., 2018). EH are widely released to surface waters mainly due to excretion by humans, medical therapy, contraceptive use and livestock-related agricultural practice (Adeel et al., 2017; Ciślak et al., 2023) and may cause adverse effects on sensitive aquatic species such as fish (Jarošová et al., 2014a, 2014b).

The strategy for water protection established by the European Union (EU) Directive 2000/60/EC (Water Framework Directive, WFD - EC, 2000; EC, 2008; EU, 2013) defines safety limit concentrations (i.e. Environmental Quality Standard (EQS)), protective for human health and the environment, for a narrow number of priority substances (PS) of EU wide concern. The current approach used for chemical water quality monitoring and to evaluate compliance with good chemical status in the EU is to compare detected chemical concentrations from targeted chemical analysis to the EQS set at european level for PS (EQS Directive, EQSD - EC, 2008) and by the Member States for the national river basin specific pollutant (RBSP). Consequently, the environmental risks posed by complex chemical mixtures and their potential additive and synergistic effects cannot be addressed. To support the identification of new PS, the Watch List mechanism was first introduced under the WFD in 2013 (EU, 2013). The EH, estrone (E1), 17β-estradiol (E2), and 17α-ethinylestradiol (EE2) were listed on the first Watch List published in 2015 (EU, 2015) and renewed in 2018 (EU, 2018) for a continuous four-year monitoring. Based on the risk assessment of the monitoring data (e.g. Loos et al., 2018) the aforementioned EH have been included in the list of PS in the proposal to amend the WFD (EC, 2022). Although these substances are the main drivers of estrogenic activity in surface water, other EDCs may contribute to the overall estrogenic activity in the aquatic environment (Gonsioroski et al., 2020; Kunz et al., 2017). Five substances classified in EU or under assessment as endocrine disruptors according to the European Chemicals Agency (ECHA) and with potential estrogenic effects are included as PS in the WFD, i.e., di (2-ethylhexyl) phthalate (DEHP), diuron, 4-nonylphenol (4-NP), 4-tert-octylphenol, terbutryn, while two have been added to the proposal of the amendment of the WFD, including bisphenol A (BPA) and triclosan. Nevertheless, other estrogenic EDCs not included in the WFD or in the proposal of the amendment of the WFD, such as tamoxifen, may also occur in the acquatic environment.

This emphasises the need for monitoring methods that account for the overall impact of chemical mixtures and, accordingly, the establishment of threshold values to assess their risk (Wernersson et al., 2015; Altenburger et al., 2019; Drakvik et al., 2020). *In vitro* effect-based methods (EBM) offer a rapid, sensitive, and cost-effective way to screen the overall compounds sharing modes of action (MoAs). Thus, they are proposed in the WFD amendment and recently approved by the Council (EC, 2022) for assessing cumulative estrogenic effects in surface waters. However, EBM validation and harmonisation are essential before policy implementation. Several *in vitro* bioassays or EBMs are included in the Organisation for Economic Co-operation and Development (OECD) test guidelines (e.g., OECD TG455) for assessing the (anti) estrogenic potency of chemicals. There are also International Organization for Standardization (ISO) standards 19040 for determining the estrogenic potential of water such as ISO Standard (2018a,b,c). Nevertheless, *in vitro* EBM responses do not automatically indicate water toxicity or unacceptable chemical quality. Thus, establishing EBM-related threshold values is crucial. Effect-based trigger values (EBTs) help distinguish acceptable from unacceptable EBM responses (Escher et al., 2021; Neale et al., 2023a). MoA-specific EBTs could be used as a guidance as follows: if the bioanalytical-equivalent concentration (BEQ) is lower than the EBT, there is no indication of risk, and if the BEQ is equal or higher than the EBT, there is indication of potential

risk and further action is warranted.

There is no single, universally accepted method for deriving EBTs. Instead, the literature presents different approaches (Leusch et al., 2014; Tang and Escher, 2014; Jarošová et al., 2014a; Kunz et al., 2015; van der Oost et al., 2017; Besselink et al., 2017; Escher et al., 2018; Brion et al., 2019; Escher et al., 2021; Neale et al., 2023a), each with its own strengths and limitations (Simon et al., 2022b). This diversity in methodologies reflects the complexity of assessing the combined effects of multiple pollutants and the need for tailored solutions depending on specific environmental contexts and objectives.

Whithin this framework, the Joint Research Centre (JRC) launched and led a joint strategic action plan for the identification of EBT values to determine the environmental risk of chemical mixtures through this interlaboratory study focused on estrogenicity EBMs. To establish the EBT, two approaches from the literature were selected: the first one (option 1) links the trigger value to the EQS according to Escher et al. (2018), while the second approach (option 2) relies on a correlation between *in vitro* and *in vivo* methods as determined by Brion et al. (2019) to demonstrate the ecological relevance of the *in vitro* methods used and the threshold set. Additionally, a third approach (option 3) was included as average of eleven bioassay-specific EBTs of option 1.

This interlaboratory study aimed at providing the basis for proposing an EBT for estrogenicity EBMs and to compare the results obatained with different EBMs to chemical analysis. To achieve this goal, eight environmental Water Samples (WS) were collected, and mixture controls either of EH (E1, E2 and EE2), or estrogenic EDCs (BPA, DEHP, diuron, 4-NP, 4-tert-octylphenol, tamoxifen, terbutryn and triclosan) were prepared. These samples were distributed and measured by the participating laboratories using EBMs (14 participants, 12 different bioassays) and/or chemical analysis (4 participants).

The results of these analyses were compared in terms of risk quotient (RQ) to determine: (i) whether the EBMs were able to identify the presence of estrogenic substances at the same level of chemical analysis; (ii) the sensitivity of the tested EBM benchmarked against the results of the chemical analysis; (iii) if the EBMs were able to detect estrogenic effects of other EDCs when EH were not present; and (iv) to provide an evidence-based ground for discussion and recommendations to the Member States on EBM applications to support the monitoring of EH and EDCs under the WFD.

2. Materials and methods

2.1. Participating laboratories

Eighteen laboratories from 11 countries including EU Member States (Austria, Belgium, Czech Republic, Finland, France, Germany, Italy, Spain, Sweden), Switzerland, and the United States of America (USA) voluntereed to participate in the interlaboratory study. Fourteen laboratories carried out EBM analysis using twelve different EBMs covering estrogenicity, while four laboratories carried out chemical analysis (Supplementary Information, SI, Table S1).

2.2. Water samples and controls

WS were collected from eight different locations near the JRC, in Ispra (Italy). The sampling points (SP) were selected based on the expected presence of EH (E1, E2 and EE2) and other estrogenic EDCs (diuron, nonlylphenol, 4-tert-octylphenol, terbutryn) according to the monitoring data collected previously by the Regional Environmental Protection Agency (ARPA, Lombardy and Piedmont, Italy) and Water Information System for Europe (WISE) database (European Environmental Agency, EEA, 2015-2022). A description of the sampling sites is summarised in Table S2 (SI). A solvent control and two artificial mixture controls, one EH mixture (EM) and one estrogenic EDCs mixture (EDCM) were included. EM and EDCM controls were prepared at a concentration equal to 1 x EQS according to Table 1. In this study, the terms EH and estrogenic EDCs will refer only to the estrogenic substances present in the EM and EDCM controls that are listed in Table 1 and SI (Section 1.1).

Table 1
Estrogenic hormones (EH) and other estrogenic endocrine disrupting chemicals (EDC) analysed using chemical methods during the interlaboratory study and used to prepare the control mixture. In blue the EH present in EM control and in white the estrogenic EDCs present in EDCM control. PS, Priority Substances; UA, under assessment (ECHA's ED assessment list). Yes, officially recognised in the EU as ED (ECHA). SERM (selective estrogen receptor modulator).

Substance	Confirmed ED with estrogenic activity	Legislation	EQS (ng/L)	Controls
Estrone (E1)	Yes	Proposal as PS candidate[a]	0.36	EM
17β-estradiol (E2)	Yes	Proposal as PS candidate[a]	0.18	
17α-ethinylestradiol (EE2)	Yes	Proposal as PS candidate[a]	0.017	
Bisphenol A (BPA)	Yes	Proposal as PS candidate[a]	10[b]	EDCM
Di(2-ethylhexyl) phthalate (DEHP)	Yes	WFD	1300	
Diuron	UA	WFD	49	
4-Nonylphenol	Yes	WFD	37	
4-tert-Octylphenol	Yes	WFD	100	
Tamoxifen	SERM	RIVM[c]	67	
Terbutryn	UA	WFD	65	
Triclosan	UA	Proposal as PS candidate[a]	20[a]	

[a] Candidate Priority Substance in the proposal WFD review and listed in the Environmental Quality Standard Directive (EQSD) review.

[b] The current EQS in the WFD proposal (EC, 2022) is much lower (0.17 ng/L) than the value selected by the JRC for the interlaboratory study (10 ng/L). During a meeting held on June 29, 2023 the experts agreed to use a concentration which would resemble the environmental concentration. The selected value corresponds to 10-folds the blank values, which are often around 1 to 2 ng/L, depending on methods used.

[c] The Dutch National Institute for Public Health and the Environment (RIVM), Moermond et al. (2018).

2.2.1. Solid-phase extraction (SPE) for chemical analysis and EBM analysis

WS and controls were loaded onto SPE cartridges and eluted following a standard operating procedure (SOP). Cartridges were kept at 4 °C until they were distributed to the participating laboratories where they were stored at −20 °C until elution. The SOP followed recommendations given by a large SPE optimisation study specifically targeting EDCs (Schulze et al., 2024). More details are in the SI (Section 1.1).

2.3. Chemical analysis

Three laboratories (#15, #17 and #18) performed the chemical analyses of EH for the eight WS extracts and controls (solvent and blank). In addition, two laboratories (#16 and #17) analysed the eight EDCs for WS extracts and controls, while laboratory #18 measured only BPA.

Although more laboratories initially performed measurements, only two laboratories and three distinct analytical methods were retained in the evaluation, as some did not follow the SOP for the elution step and used their own protocols instead (see SI, Fig. S1). The quantification of EH (E1, E2, and EE2), BPA, diuron, terbutryn, and tamoxifen was performed using liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS), with the latter three compounds analysed using a

dedicated LC-MS/MS method. The simultaneous determination of BPA, DEHP, 4-NP, 4-tert-octylphenol, and triclosan was carried out using gas chromatography coupled with tandem mass spectrometry (GC-MS/MS). More details are in the SI (Section 1.2 and Chapter 2).

2.4. Effect-based methods

Fourteen laboratories analysed the WS and controls using EBM. Expert laboratories were allowed to choose estrogenicity bioassays according to their preferences and expertise without limitations to a specific method (i.e., standardised bioassays, commercial kits and methods described in scientific literature were allowed) providing an overall EBM pool which included as many different bioassays as possible for the study (SI, Table S1). Twelve bioassays were included in the study, comprising a range of cell-based, yeast-based, and protein-binding assays (Table 2, and SI Section 1.3).

Biological activities of the controls and WS extracts in the bioassays were expressed as BEQ ($ng_{positive\ control\ equivalent}/L$), representing the concentration of the positive control that elicits the same effect as the sample. The natural hormone E2 was used as a reference compound to determine the estrogenic potential of the samples and controls, which were expressed as E2-equivalent concentration (EEQ, ng_{E2}/L). More details are in the SI (Section 1.3.4).

2.5. Effect-based trigger values

A risk assessment based on two existing datasets from Könemann et al. (2018) and Gómez et al. (2021) was performed using two EBT derivation options based on Escher et al. (2018) and Brion et al. (2019), with the aim to compare results from EBMs and chemical analyses (JRC, unpublished study). The outcome of the two case studies indicated that the chosen EBT values were suitable for risk assessment of WS. Consequently, we selected these approaches as a starting point for deriving estrogenicity EBT values (Table 3). The first one (option 1) links the trigger value to environmental quality standards (EQS) according to the EBT derivation option B described by Escher et al. (2018) limited to E1, E2 and EE2. The EBT values calculated based on this approach and the updated annual average EQS (AA-EQS) were derived for all bioassays. The second one (option 2) is based on a correlation between *in vitro* and *in vivo* methods determined by Brion et al. (2019). To obtain intermediate values between the EBT simulated by option 1 and the EBT determined by Brion et al. (2019), an assessment factor (AF) of 2 was applied. Since Option 1 is based on EQS (derived with AF) we applied a low AF to harmonise the approaches. The EBT values determined by Brion et al. (2019) divided by 2 were used for four *in vitro* bioassays (Table 3). The third one (option 3) is determined as average of all EBTs in option 1. Table 3 shows the three different EBT options used in this study. More details can be found in the SI (SI, Chapter 3 and Table S32).

Table 2
Bioassays tested in the interlaboratory study.

Cell-based transactivation assays	Yeast-based transactivation assays	Protein-binding assay
ERα-CALUX	Yeast Estrogen Screen (YES)	Ligand Binding Estrogen Receptor Assay (LiBERA)
MELN assay	YES XenoScreen	
ERα-GeneBLAzer	Arxula YES	
T47D-Signosis	Lyticase-assisted YES (L-YES)	
VM7Luc ER transactivation (TA) assay	Planar YES (p-YES)	
Estrogen receptor bioassay for environmental sample monitoring (INDIGO Biosciences)		

2.6. Risk quotients (RQs)

The individual chemical RQs were calculated for each EH (E1, E2 and EE2) and EDC (BPA, DEHP, diuron, 4-NP, 4-tert-octylphenol, tamoxifen, terbutryn and triclosan) at all SPs and controls (EM and EDCM) by dividing the measured concentration by the respective AA-EQS. Then, the cumulative risk was assessed by $\sum$RQs (i.e. summing the individual RQs of E1, E2 and EE2) indicating a risk if $\sum RQs \geq 1$ and a low or no risk when $\sum RQs < 1$.

Regarding the EBM RQ, the EEQ determined by the participating laboratories through *in vitro* EBMs were divided by each of the three EBT options shown in Table 3, giving values ≥ 1 that inform about risk, or <1 which indicate a low or no risk.

Concerning non-detects, three data scenarios were considered: *i)* worst-case when concentrations below the limits of quantification (LOQ) were substituted with LOQ, *ii)* common approach when concentrations below LOQ were assumed equal to LOQ/2, and *iii)* lower bound when concentrations below LOQ were set as zero. The worst-case scenario was used in this study for the RQ calculation for both chemical and EBM analyses. For RQ of EH, also the estimates based on spiking were used (SI, sections 2.4 and 2.5).

The compliance or non-compliance with the chemical analyses of EH, EDCs or the sum of EH and EDCs were examined for each sampling point. Compliance (conformity or accordance) indicates that the results of the EBM analysis align with the chemical analysis (reference method). Compliance is defined by categorical agreement, which is achieved when both methods classify a site as either risk ($RQ \geq 1$) or non-risk ($RQ < 1$). Non-compliance occurs when the two methods provide opposite classifications. Borderline values shall be rounded to the nearest decimal; for example, a value of 0.95 shall be rounded to 1.00. EBMs need to demonstrate an equal sensitivity compared to chemical analysis to ensure the same safety level.

2.7. Data analysis

Data were analysed in terms of RQs. The RQs based on EBM results estimated by different bioassays or participating laboratories were evaluated using firstly ANOVA test for statistical significance of differences and additionally by several correlation tests (goodness of fit, Pearson and Spearman).

ANOVA checks whether the group means of RQs (in our analysis the various bioassays or laboratories) are different from the overall mean of the data by checking the variance of each individual group against the overall variance of the data. If one or more groups are outside the range of variation by the null hypothesis (all group means are equal), then the test of differences is statistically significant. The ANOVA output tables are included in the SI (Chapter 5).

The correlation of the EBM estimates with the chemical RQs was assessed in three manners. Firstly, by the coefficient of determination (R^2) representing the goodness of fit, a statistical measure of how well the regression line approximates the actual data. Thereafter, statistical correlation tests of Pearson (a linear association between two normally distributed random variables) and Spearman (more robust to outliers and without distributional assumptions comparing to Pearson) were applied. According to the commonly accepted way to interpret R^2 scores, the values near one indicate a high/strong correlation and opposite – the scores near zero suggest weak or no correlation. Similarly, the absolute values of Pearson and Spearman correlation coefficients are interpreted as follows: strong ≥ 0.7, good in the interval $0.5 \leq - < 0.7$, moderate in the range $0.3 \leq - < 0.5$, low < 0.3, no correlation $= 0$. For the interpretation, a preference was given to the high Spearman correlations but the decisions were supported also by R^2 and Spearman evaluation (details in SI, Section 5.5.1).

Table 3
Effect-based trigger value (EBTs) options chosen and evaluated during this interlaboratory study, the values are expressed as ng E2-equivalents/L. The relative potency (REP) used to determine the EBT options 1 and 3 can be checked in the SI (Table S31). NA, not available.

EBT option	Source	ERα-CALUX	MELN	ERα-Gene-BLAzer	T47D-Signosis	VM7Luc4E2 TA	INDIGO	YES	A-YES	L-YES	pYES	LiBERA
Option 1	JRC derived following Escher et al. (2018) option B	0.073 [a]	0.094 [a] [b]	0.079 [a], [c]	0.151 [d]	0.068 [e]	0.093 [f]	0.095 [g]	0.087 [h]	0.084 [i]	0.079 [c]	0.106 [j]
Option 2	EBT as half value determined by Brion et al. (2019)	0.14	0.28	0.12	NA	NA	NA	NA	NA	NA	0.25	NA
Option 3	Average of option 1	0.092										

[a] REP as average of the values obtained by Könemann et al. (2018), Simon et al. (2022b).
[b] REP as average of the values reported by Neale et al. (2018), Könemann et al. (2018) and Serra et al. (2020).
[c] REP obtained from Könemann et al. (2018).
[d] REP obtained from Laboratory #11.
[e] REP obtained from Test Guideline No. 455, OECD (2021).
[f] REP calculated as average REP value from the other cell-based bioassays (ERα-CALUX, MELN, ERα-Gene-BLAzer, T47D-Signosis, VM7Luc4E2 TA) included in this study.
[g] REP determined as average of literature values from Jarošová et al. (2014a), Murk et al. (2002), Rutishauser et al. (2004), Schultis et Metzger, 2004, Leusch et al., 2010 (SI), Zhao et al. (2011) (SI).
[h] REP obtained as average of Simon et al. (2022b) and laboratory #9.
[i] REP obtained from Simon et al. (2022b).
[j] REP derived by the JRC.

3. Results

3.1. Results of the chemical analysis

3.1.1. Chemical analysis of estrogenic hormones (EH)

The results obtained by the laboratories #15, #17 and #18 are shown in the SI (Chapter 2). Participating laboratory #15 reported lower concentrations for the EM control respect to the spiked ones. This laboratory quantified the EH in several SPs (2, 4, 5, 6 and 8), while participating laboratory #17 was able to quantify the considered EH only in SP3 and SP4 with a low level of recovery. For the EM control, the latter reported non-detected concentrations (below the limit of detection, LOD = 0.02 ng/L), Laboratories #15 and #17 did not apply the elution protocol included in the SOP circulated by the JRC which recommends two solvents - methanol and ethyl acetate. Instead, they used their own elution protocols based only on methanol. This could explain the observed difference between reported analytical results, including low recoveries in some cases. Therefore, after a quality check as shown in the workflow (SI, Fig. S1), it was decided to exclude the results of laboratories #15 and #17, but to keep these of laboratory #18. Further analyses and comparisons are presented (in the SI, Chapter 2, Sections 2.1, 2.3 and 2.5).

Fig. 1 visualises the expected cumulative chemical risk of EH estimated by summing RQs of the individual substances, according to the results of laboratory #18 for the EM control and the 8 SPs. Estimates based on spiking were used for E2 and EE2 in the cases the values were below the LOQ (SI, section 2.4). All SPs showed $\sum$RQs>1, i.e. a presence of cumulative estrogenic risk. In addition, the figure evidences a variability of main drivers of the cumulative estrogenic risk at the different sampling locations. Indeed, E1 is the main driver at SPs 2, 3, 5 and 6, while for the EM control and SPs 1, 4, 7 and 8 this is the combined action of E2 and EE2.

The driving role of individual substances in the cumulative risk of the considered EH is confirmed also by statistical analysis with ANOVA test

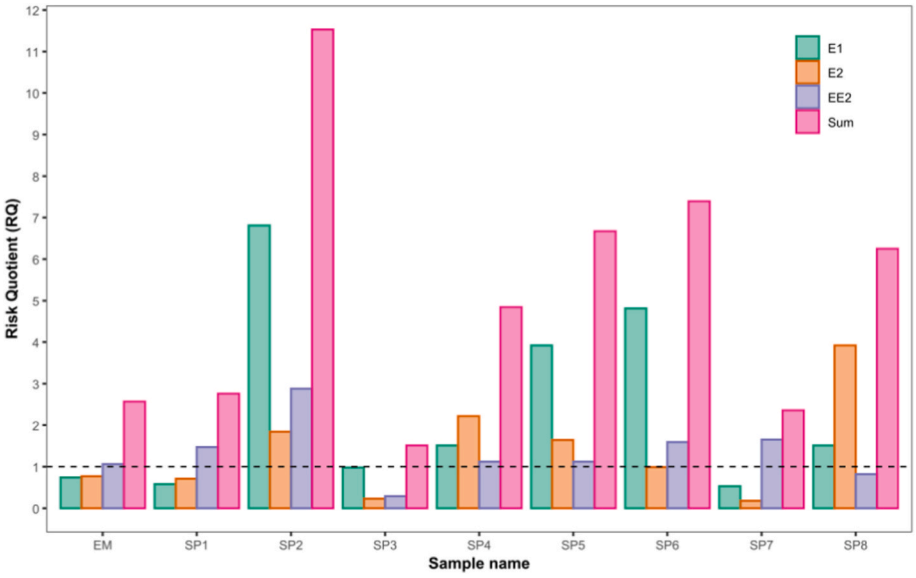

Fig. 1. Cumulative chemical risk of estrogenic hormones (E1, E2 and EE2) estimated by summing the RQs of the individual substances, according to the analytical results of laboratory #18, for the control and the 8 sampling points (SP). The dashed line marks the critical threshold of RQ = 1, above which there is indication of risk.

(SI, Section 5.1.1). Generally, E1 is dominating or contributing to the cumulative estrogenic risk.

3.1.2. Chemical analysis of further estrogenic EDCs

Two laboratories (#16 and #17) reported results for the chemical analyses of the eight EDCs for WS extracts and controls, while laboratory #18 measured only BPA. For the reasons mentioned in section 3.1.1, the results from laboratory #17 were excluded from the analysis (see SI Chapter 4, Sections 2.2, 2.3, 2.4, 2.5). Fig. 2 presents the worst case scenario (non quantified = LOQ) of the cumulative chemical risk posed by the measured EDCs, estimated by summing the RQs (concentration in sample divided by EQS) of the individual compounds, according to the results of laboratory #16 and laboratory #18 (only BPA). A presence of cumulative risk ($\sum$RQs>1) was found at a majority of the SPs (2, 3, 4, 5, 6 and 7) and it was observed that BPA, which was quantified at all SPs except the first one, is a driver of this risk. The remaining SPs (1 and 8) showed no presence of cumulative risk in regard to the eight EDCs analysed in this study. The dominant role of BPA as a driver of the cumulative risk, amongst the considered EDCs, is confirmed also by statistical analysis with ANOVA test (SI, Section 5.1.2).

3.2. Results of the EBM analyses

The EBM results were reported by the participating laboratories as EEQ. EEQ values by laboratory and bioassay are reported in the SI (Chapter 2 and section 5.4.1, Tables S33 and S34). In the following sections the EBM results are discussed as RQ determined using the three EBT options evaluated in this study (see Table 3, Section 2.5).

3.2.1. Analysis of EBM results using different EBT options

The impact of using the different EBT options for calculating the RQ on the EBM results is described in the SI (Section 5.2). The three EBT options were available for four bioassays (ERα-CALUX, MELN, ERα-GeneBLAzer and p-YES). For the rest of bioassays (VM7Luc4E2, T47D-Signosis, INDIGO, YES, A-YES, L-YES and LiBERA) only option 1 and option 3 were available.

EBT option 1 was taken as reference to compare the results obtained using option 2 and option 3. EBT option 2 would reduce the RQ by more than 30%, whereas option 3 would result in changes below 20%, except for the bioassays based on VM7Luc4E2 (−26%, underestimation of risk) and T47D-Signosis (+64% overestimation of risk) cell lines.

3.2.2. Analysis of EBM results of participanting laboratories using the same or variants of similar bioassays

Some participating laboratories used the same or variants of similar bioassays. For instance, ERα-CALUX was applied by three laboratories (#1, #7 and #12) and ERα-GeneBLAzer by two laboratories (#6 and #13), while variants of YES bioassay were performed by six laboratories (#2, #3, #4, #8, #9 and #12). Among them two laboratories used A-YES (#4 and 9#). All WS and controls analysed by laboratory #3 using YES Xenoscreen were below the LOQ. This bioassay had an LOD much higher (4.9 ng/L) than the EBTs evaluated in this study, therefore, it was excluded from further analysis. Among the two laboratories using ERα-GeneBLAzer (#6 and #13), laboratory #13 reported quantified EBM values only for two SPs, while laboratory #6 reported effect data for all SPs. A possible explanation for these differences is the sample stability in the cartridges that could have been compromised due to a long shipment time. For this reason, ERα-GeneBLAzer results from laboratory #13 were not averaged with those of laboratory #6 and excluded from further analyses.

A detailed analysis of the differences between EBM results of laboratories performing the same or similar bioassays (i.e., ERα-CALUX and A-YES) is presented in the SI (Section 5.3). In this study, the results of the EBM are shown as average in case of laboratories performing the same bioassay. In the case of ERα-CALUX, some differences were observed between laboratories with laboratory #7 reporting lower values. This

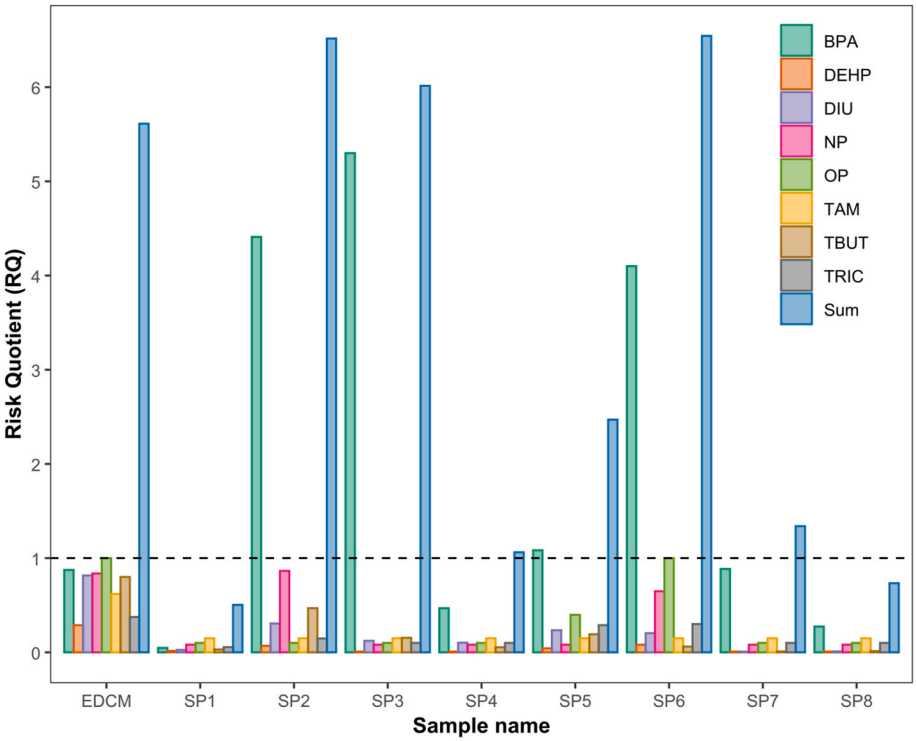

Fig. 2. Cumulative chemical risk of EDCs estimated by summing RQs of the individual compounds, according to the analytical results of laboratory #16 and #18 (only for BPA), for the EDCM control and the 8 sampling points. BPA: bisphenol A; DEHP: di(2-ethylhexyl) phthalate; DIU: diuron; NP: 4-nonylphenol, OP: 4-tert-octylphenol; TAM: tamoxifen; TBUT: terbutryn; 4-nonylphenol; TRIC: triclosan. The dashed line marks the critical threshold of RQ = 1, above which there is indication of risk.

may be due to the delay between sample shipment, extraction and analysis. Laboratory #2 repeated the analysis of WS extracts after two weeks stored at −80 °C and observed a significant decrease in the estrogenic activity indicating that sample stability may be compromised by the time elapsed between SPE and analysis.

3.2.3. Overview of the EBM RQs using EBT options 1, 2 and 3

EBM RQ results obtained for the controls (EM and EDCM) and the eight WS (SP1-SP8) estimated using EBT options 1, 2 and 3 are shown in Fig. 3 and in the SI (Section 5.4).

INDIGO and T47D-Signosis showed generally higher mean RQ values compared to the other bioassays when using EBT option 1 or option 3. ANOVA statistics was performed to check about significant or non-significant differences between means of RQ estimated by EBM results of the different bioassays and participating laboratories (SI, Section 5.4.2.). ANOVA tests confirmed the observations concluded by the descriptive statistics of RQs for EBT options 1 and 3. For EBT option 2. ANOVA tests showed differences between ERα-GeneBLAzer and the other bioassays.

3.3. Comparison of the EBM RQs with the cumulative chemical RQs

The EBM RQs determined using the three EBT options were compared to the chemical analysis of EH and EDCs showing very similar trends. While the choice of EBT option may increase or decrease the RQs, the direction and pattern of the fitting lines remain consistent across all options. Consequently, only the correlation results corresponding to the EBT option 1 are shown. The correlation analysis was performed as described in the materials and methods (section 2.7) and includes information about the coefficient of determination as well as Pearson and Spearman tests. Detailed information is presented in the SI (Section 5.5.1).

3.3.1. Overall correlation between EBM and chemical analysis

The main findings of the correlation analysis between EBM results and chemical analyses of the three EH are presented in Fig. 4. Overall, four EBMs, ERα-CALUX, MELN, T47D-Signosis and L-YES performed well, showing high coefficient of determination and statistically significant strong Spearman correlations. ERα-GeneBLazer, VM7Luc4E2, INDIGO, YES, p-YES and LiBERA showed good while A-YES weaker correlations.

The correlation of EBM results was also analysed considering only the EDCs included in this study and the mixture of all substances (EH and EDCs). Briefly, LiBERA and ERα-GeneBLazer results showed strong and good correlations (respectively) to the chemical RQs of the estrogenic EDCs; MELN, T47D-Signosis, L-YES and p-YES had moderate to low correlations, while the rest of the bioassays showed no correlations. The comparison with the mixture of all substances (EH and estrogenic EDCs) indicated strong/good correlations for ERα-CALUX, MELN, ERα-GeneBLAzer, T47D-Signosis, L-YES, p-YES and LiBERA. Complete results are presented in the SI (Sections 5.5.1 and 5.5.2).

3.3.2. Analysis of the results by sampling point

Table 4 summarises the compliance assessment of bioassays against cumulative chemical risk estimated for EH, EDCs, or their sum at each SP. Compliance was also evaluated for controls (EM and EDCM, see SI, section 5.5.4). Fig. 5 visualises the RQs estimated by chemical analysis (EH, EDCs and EH + EDCs) and different EBMs at each SP under EBT option 1, along with deviations as percentages. Compliance/non-compliance results are summarised in Table 5 and displayed in Fig. 6 as compliance percentages per EBT option.

Table 4 shows that ERα-CALUX (option 1), MELN (options 1 and 3), ERα-GeneBLAzer, T47D-Signosis, VM7Luc4E2, INDIGO, YES and p-YES are compliant with cumulative chemical risk of EH at all SPs.

A-YES did not indicate cumulative chemical risk of EH when there was no individual risk of E1 and did not respond to cumulative risks arising exclusively from EDCs when the overall cumulative risk of EH was low (SP3).

ERα-CALUX (option 3), MELN (option 2), L-YES could not identify estrogenic chemical risk at only one site (SP7) under condition of low level of cumulative risk created either individually only by EE2 (missing individual risk from E1) or cumulatively by EDCs.

ERα-CALUX (option 2) did not identify the estrogenic chemical risk at two sampling sites (SP3 and SP7), where there is cumulative risk of EDCs but the cumulative risk of EH is not high.

LiBERA did not identify estrogenic chemical risk at two sampling sites (SP1 and SP7), at these sites the presence of cumulative risk of EH is dominated by EE2 and no individual risk of E1 and E2. Moreover, there is no cumulative risk of estrogenic EDCs at SP1, while at SP7 the cumulative risk of EDCs is slightly above 1 and dominated by BPA.

The compliance/non-compliance checks of bioassays comparing to cumulative risk estimated by chemical analysis of only estrogenic EDCs showed non-compliance for SP1 and SP8. There is no cumulative risk of EDCs at these SPs, but most bioassays indicated risk, which was triggered by EH.

The EBM RQs showing higher overall correlation with chemical analysis and higher rate of compliance at sampling points are usually near to or in the ±50% range of the cumulative chemical RQs (Fig. 5).

As presented in Table 5, ERα-CALUX (option 1), MELN (options 1 and 3) ERα-GeneBLAzer, T47D-Signosis, VM7Luc4E2, INDIGO, YES, and p-YES showed 100% compliance with chemical risk estimates, while others, A-YES, ERα-CALUX (option 2) and LiBERA, exhibited non-compliance at 12.5–25% of sites. Non-compliance primarily occurred when cumulative risk was driven solely by estrogenic EDCs or low-level EH (e.g., E1). Additional analysis for the mixture of EH and EDCs is provided in the SI (Sections 5.5.4 and 4.5.5).

Using EBT option 1 and when chemical analysis of only EH was considered (Fig. 6, A), 8 out of 11 bioassays (73% of all bioassays) achieved full compliance (100%); for EBT option 2, 2 out of 4 bioassays (50% of those for which option 2 is available) achieved full compliance; and for EBT option 3, 7 out of 11 bioassays (64% of all bioassays) achieved full compliance.

Comparing the results of the EBM RQ with the chemical analysis of only estrogenic EDCs (Fig. 6, B), 9 out of 11 bioassays (82%) reach the 75% compliance when EBT option 1 is used, 2 out of 4 (50%), when EBT option 2 is used and 8 out of 11 (73%) when EBT option 3 is used.

4. Discussion

This study confirmed that effect-based monitoring provides significant advantages for assessing water quality regarding EH and estrogenic EDCs, aligning with previous research (Carvalho et al., 2014; Gómez et al., 2021; Simon et al., 2022a; Glineur et al., 2024). EBM capture in a single measurement a broader spectrum of chemicals contributing to biological effects, beyond what chemical analysis alone can quantify. Results also showed that both EH and EDCs exert estrogenic activity, effectively detected by various estrogenicity EBMs. Notably, in three samples (SP1, SP4, SP5 and SP8) estrogenic activity observed in bioassays could not be explained only by EH and EDCs (SI, sections 5.5.7 and 5.5.8).

The study also identified possible key factors – sample pre-treatment, elution protocols, storage conditions, data evaluation methods, type of EBM used, and EBT derivation approach – that are critical for enabling the use of EBMs. The EBM analysis of surface WS usually requires a sample enrichment step. In the current study, SPE was used to concentrate the samples by a factor of one thousand. Blank and solvent controls are essential to determine whether estrogenic effects were triggered by impurities captured during the extraction process, due to the SPE materials and solvents used for WS enrichment (Neale et al., 2018). Among the EH and EDCs included in the study, DEHP and 4-NP were detected by chemical analysis in the blank and solvent controls. For 4-NP, the concentration was below the LOQ and 74-fold lower than the EQS.

Fig. 3. Changes on the RQ results using the different bioassays evaluated during this interlaboratory study by sample, when EBT options 1 (A), 2 (B) and 3 (C) are used. Please note that option 2 is available only for four bioassays. The red dashed line indicates the risk threshold (RQ = 1) above which there is indication of risk.

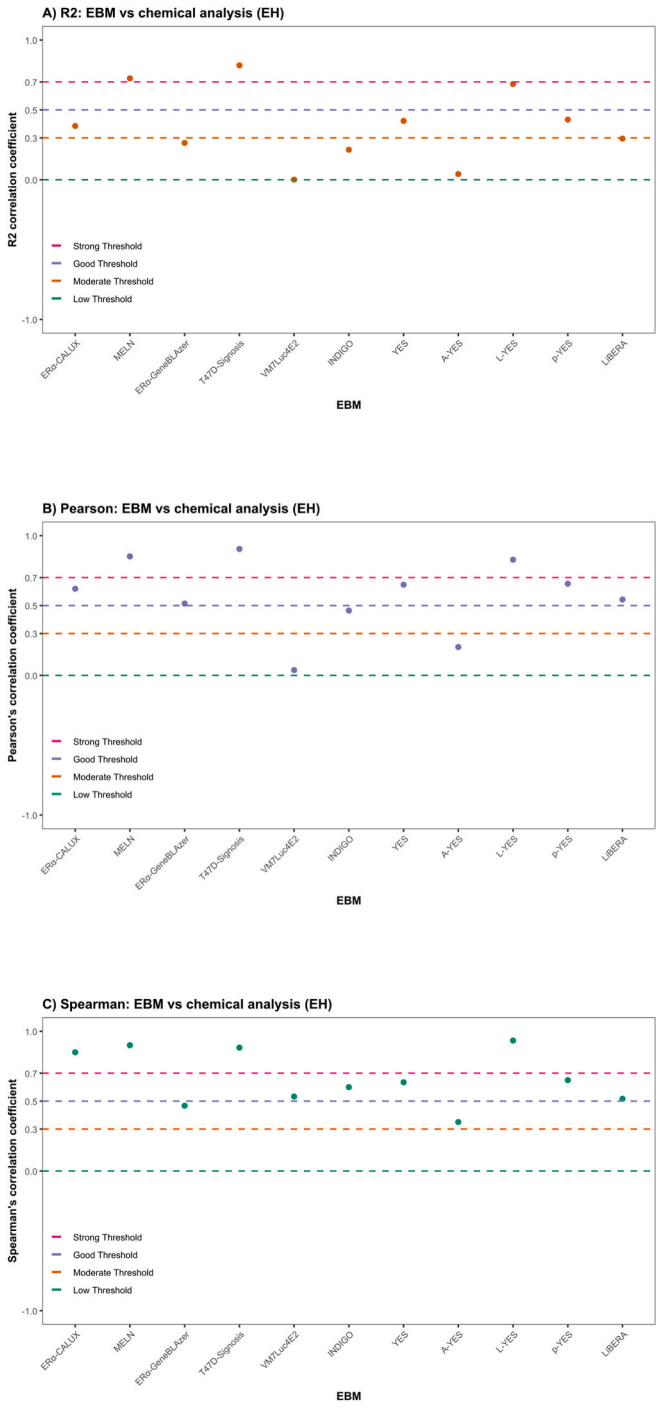

Fig. 4. Overall evaluation of EBMs by correlations of RQs estimated by EBM and cumulative chemical analysis results for the mixture of estrogenic hormones (EH). A) R^2, B) Pearson and C) Spearman correlation coefficients. Interpretation of correlations (lines in the figure): - strong ≥ 0.7, good $0.5 \leq - < 0.7$, moderate $0.3 \leq - < 0.5$, low <0.3. The strong correlations shown on the figure correspond to statistically significant results. The correlation of EBM results with chemical analysis was verified also considering only estrogenic EDCs and the mixture of all substances (EH and EDCs) (SI, Section 5.5.1).

Regarding DEHP, the value was slightly higher than the LOQ but 65-fold lower than the EQS. No effects were detected in the blank and solvent controls by most EBMs, only in few cases with a very low signal (below the LOD), showing that the contribution of the solvents and impurities to the estrogenic activity was minimal, and that there was no cross contamination between samples.

Despite the challenges of chemical analysis, some participating laboratories identified EH and selected estrogenic EDCs across eight sampling locations. However, differences among participants were observed and may be attributed to variations on the application of the SPE protocol. The detection of all substances in the EM and EDCM controls further indicates that the analytical methods employed are sufficiently sensitive to detect these compounds in surface water at the EQS level, except for EE2. The chemical analysis of triclosan was also difficult showing low recovery (40%), with a concentration in the EDCM control 2.5- fold lower than the EQS.

E1 and BPA were the primary drivers to estrogenic activity in the WS. E1, E2, EE2, BPA and triclosan are not yet under legal frame monitoring program i.e. WFD, meaning no specific measures are currently in place to control their concentrations in EU water bodies. In contrast, other EDCs included in this study are subject to regulatory measures, which may explain why their concentrations remain below the EQS and do not pose a significant risk for the aquatic environment.

Differences between laboratories testing the same EBM were observed (e.g., ERα-CALUX and ERα-GeneBLAzer) which could be attributed to a compromised stability of the sample due to sample processing delay and shipment conditions. Sample stability is influenced by the time elapsed between sample collection, shipment, preparation of the sample extract, and analysis. Additionally, variations in elution parameters, such as reagents and instruments, could also contribute to these differences. These findings highlight the importance of robust methods and SOP for WS collection, processing and analysis. The repeatability of ISO bioassays in this study was assessed by averaging results from laboratories using the same bioassay, even though some differences were observed between them. The uncertainty in the EBM results is also influenced by the estimation of EEQ, highlighting the importance of accurate input data and assumptions in the overall analysis (Wagner et al., 2013).

Apart from the variations in the sample processing and analysis of results, differences on the EBM responsiveness might be attributable to the intrinsic nature of the bioassays, which are based on different kinds of biological systems i.e., yeast, human cell lines, other mammalian cell lines, recombinant proteins and also the type of estrogen receptor (α, β) expressed (Könemann et al., 2018; Gómez et al., 2021; Simon et al., 2022b).

The EBT derivation approach is also critical for assessing EBM results. EBTs are essential for water quality effect-based monitoring as discussed in the works of Escher et al. (2018, 2021), Simon et al. (2022b), and Neale et al. (2023a).

Here we explored three EBT derivation options (options 1, 2 and 3) and our findings are consistent with previous research showing that EBM results depend on the EBT option applied. The bioassay specific-EBT derivation (option 1) showed to be the most protective for most bioassays with eight out of eleven bioassays included in the study reaching the 100% compliance with chemical analysis of EH and EDCs in WS. The use of EBT option 2 reduced the RQ by more than 30% leading to non-compliance for the EM and EDCM controls and for two water samples. The use of EBT option 3 (0.09 ng E2-Eq/L) resulted in rchanges below 20%, except for two bioassays, and showed to be slightly less protective compared to option 1, with seven bioassays out of eleven reaching the 100% compliance. Within this dataset and under the stated assumptions (e.g. worst-case scenario, chemical data exclusion after quality check), option 1 resulted the most appropriate methods because it covers all bioassays included in the study and showed to be the most conservative. Therefore, the following paragraphs are referred to EBT option 1.

The EBMs applied in this study detected estrogenic activity in the natural WS. Additionally, the EM and EDCM controls were also active in most EBMs, confirming the assays' responsiveness to the target compounds. All EBMs except LiBERA and MELN showed estrogenic activity and/or risk in the EM control. While all EBMs except ERα-CALUX, L-YES, and LiBERA indicated estrogenic activity and/or risk at EDCM. These

Table 4
Check for compliance or non-compliance of EBMs using EBT options 1, 2 and 3 with cumulative risk estimated by chemical analyses at SPs regarding mixture of EH; EDCs; and the combination of EH and EDCs. Compliance refers to the degree to which the results of the EBM risk analysis align with the chemical risk analysis, non-compliance occurs when the results of the EBM risk analysis deviate from the values from the chemical risk analysis. **The level of risk is given in brackets (<1: no risk).**

Sample	Characterization of individual and cumulative chemical risk at SPs by RQs	Characterization of EBM risk by bioassay and EBT option	Compliance/Non-compliance comparing to the cumulative chemical risk of EH	Compliance/Non-compliance comparing to the cumulative chemical risk of estrogenic EDCs	Compliance/Non-compliance comparing to the cumulative chemical risk of EH and estrogenic EDCs
SP1	Cumulative risk of EH (2.8) which is dominated by EE2 (1.5). No individual risk of E1 (0.6) and E2 (0.7). No cumulative risk of EDCs.	LiBERA options 1 (0.8) and 3 (0.9) All others for options 1, 2 and 3	Non-compliance Compliance	Compliance Non-compliance	Non-compliance Compliance
SP2	High cumulative risk of EH (11.5) which is dominated by E1 (6.8) followed by EE2 (2.9) and E2 (1.8). There is also a cumulative risk of EDCs (6.5) which is dominated by BPA (4.4)	All for options 1, 2 and 3	Compliance	Compliance	Compliance
SP3	Cumulative risk of EH exists but is not high (1.5) and is dominated by E1 (1.0). No individual risk regarding E2 (0.2) and EE2 (0.3). Observed a cumulative risk of EDCs (6) which is dominated by BPA (5.3).	A-YES options 1 (0.4) and 3 (0.4) ERα-CALUX option 2(0.2) All others for options 1, 2 and 3	Non-compliance Compliance	Non-compliance Compliance	Non-compliance Compliance
SP4	Cumulative risk of EH (4.8) which is dominated by E2 (2.2) followed by E1 (1.5) and EE2 (1.1). No cumulative risk of EDCs.	All for options 1, 2 and 3	Compliance	Compliance	Compliance
SP5	Cumulative risk of EH (6.7) which is dominated by E1 (3.9) followed by E2 (1.6) and EE2 (1.1). Cumulative risk of EDCs (2.5) exists and is dominated by BPA (1.1)	All for options 1, 2 and 3	Compliance	Compliance	Compliance
SP6	Cumulative risk of EH (7.4) which is dominated by E1 (4.8) followed by EE2 (1.6) and E2 (1). Cumulative risk of EDCs (6.5) exists and is dominated by BPA (4.1.)	All for options 1, 2 and 3	Compliance	Compliance	Compliance
SP7	Cumulative risk of EH (2.4) which is dominated by EE2 (1.65). No individual risk of E1 (0.5) and E2 (0.2). Cumulative risk of EDCs exists but is not high (1.3) and is dominated by BPA (0.8).	L-YES options 1(0.2) and 3 (0.4) LiBERA options 1 (0.4) and 3 (0.5) ERα-CALUX option 2 (0.5) and 3 (0.8) MELN option 2 (0.4) All others for options 1, 2 and 3	Non-compliance Compliance	Non-compliance Compliance	Non-compliance Compliance
SP8	Cumulative risk of EH (6.3) which is dominated by E2 (3.9) followed by E1 (1.5) and EE2 (0.8). No cumulative risk of EDCs.	All for options 1, 2 and 3	Compliance	Non-compliance	Compliance

findings are consistent with those of previous studies showing that ERα-CALUX is more responsive to EH, particularly to E2 and EE2 than estrogenic EDCs compared to other EBMs (Gómez et al., 2021). Carvalho et al. (2014) and Gómez et al. (2021) found that LiBERA responded to estrogenic EDCs when EH were not present. The current comparison between LiBERA and chemical analysis of EDCs showed a strong correlation supporting the previous observation that LiBERA is more responsive to estrogenic EDCs other than EH.

Compliance between chemical analysis and EBMs for assessing estrogenic exceedances (EH + EDCs) in WS ranged from 75% to 100%. ERα-CALUX, MELN, T47D-Signosis, and L-YES showed strong correlations with EH chemical analysis, while ERα-GeneBLAzer and LiBERA correlated well with EDC analysis. Most bioassays (ERα-CALUX, MELN, ERα-GeneBLAzer, T47D-Signosis, VM7Luc4E2, INDIGO, YES and p-YES) demonstrated 100% compliance with the cumulative chemical risk of EH at all SP. Three SP (SP1, SP3 and SP7) provided key features of estrogenic risk. At SP1, the cumulative chemical risk from EH was driven by EE2 with no individual or cumulative risk of other EDCs. EBMs excluding LiBERA indicated risk. At SP3, individual hormones concentrations were below (E2, EE2) or equal (E1) to the EQS. EBMs, excluding A-YES, detected activity, suggesting that other unmeasured estrogenic EDCs can contribute to the overall estrogenic activity. Finally, SP7 showed BPA-driven EDCs risk, with EBMs (excluding LiBERA and L-YES) confirming risk. These findings highlight the value of EBMs in capturing combined estrogenic effects and untargeted EDCs, complementing traditional chemical analysis for comprehensive water quality assessment - though some bioassays require refinement. ERα-CALUX, MELN, ERα-GeneBLAzer, YES and p-YES showed good correlations with chemical RQs of EH and full compliance with the estimated cumulative risk, making them suitable for estrogenicity monitoring. While INDIGO and T47D-Signosis also aligned well with chemical data and full compliance, their results differed statistically from other bioassays, suggesting methodological or sensitivity variations. For estrogenic EDCs monitoring, only LiBERA and ERα-GeneBLAzer exhibited strong or good correlations (75% compliance) in WS analysis. ERα-CALUX showed no significant correlation and a lower compliance rate in WS analysis, indicating limited utility for estrogenic EDC monitoring when EH are absent.

Refinement actions are suggested based on the outcome of this interlaboratory study when differences among results from the EBM and chemical analysis are identified (e.g. correct application of the SOP, minimising time lapse between sample collection, extraction and analysis). In case the differences mainly resulted from the intrinsic nature of the bioassays, the application of an adjustment factor is being evaluated, as exemplified by LiBERA. To enhance its sensitivity, Ferrero et al. (2014) modulated the binding properties of the ERαLBD by rationally modifying its amino acid composition, resulting in various variants. Among these, the M421F-ERα LBD variant exhibited 2- to 6-fold

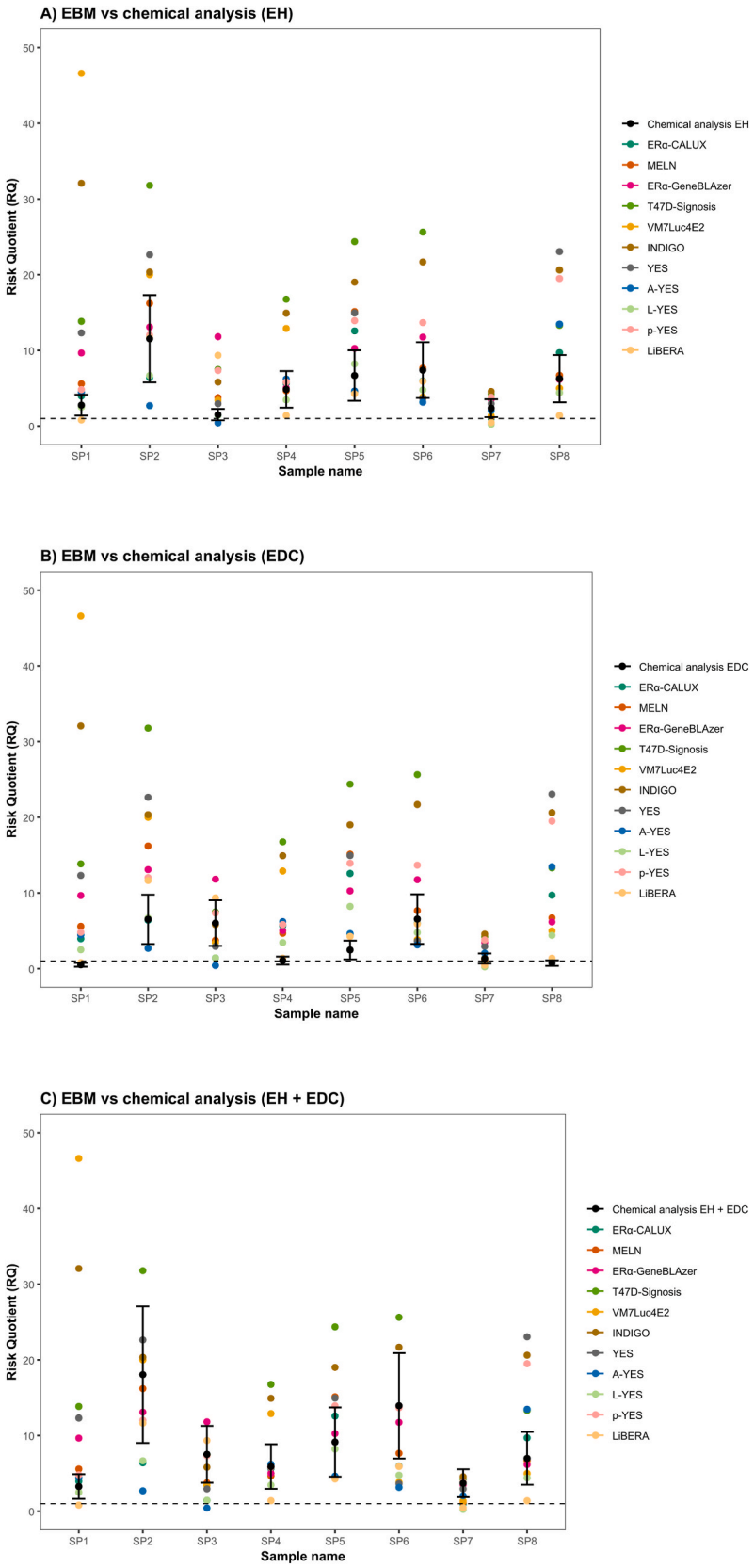

Fig. 5. Comparison of RQs estimated by different EBMs and chemical analysis for A) estrogenic hormones (EH), B) estrogenic EDCs and C) EH and EDCs at each sampling location (SP1-SP8) for EBT option 1 (for options 2 and 3 the graphs can be seen in the SI, section 5.5.3). The error bars represent ±50% of the cumulative chemical RQs. The dashed line marks the critical threshold of RQ = 1, above which there is indication of risk.

Table 5
Summary of compliance/non-compliance analysis of bioassays relative to the cumulative chemical risk estimated by chemical analyses (∑RQs >1) for EH and EDCs at eight sampling locations (SP1-SP8). EBT options are specified in parenthesis when results differ across options.

Bioassay	Compliance/Non-compliance	Notes
ERα-CALUX (option 1) MELN (options 1 and 3) ERα-GeneBLAzer T47D-Signosis VM7Luc4E2 INDIGO YES p-YES	Compliance at 100% of sites.	ERα-CALUX (option 1), MELN (options 1 and 3), ERα-GeneBLAzer, T47D-Signosis, VM7Luc4E2, INDIGO, YES, and p-YES showed compliance with cumulative chemical risk of EH at all sampling points. The cumulative risk from EH was complemented at 5 locations from the cumulative risk generated by the EDCs.
ERα-CALUX (option 3) MELN (option 2) L-YES	Compliance at 87.5% of sites. Non-compliance at 12.5% of sites.	ERα-CALUX (option 3), MELN (option 2) and L-YES could not identify estrogenic chemical threats at lower levels of cumulative risk created either individually only by EE2 (missing individual risk from E1) or cumulatively by EDCs.
A-YES	Compliance at 87.5% of sites. Non-compliance 12.5% of sites	A-YES could not forecast cumulative chemical risk of EH when no individual risk of E1, E2 or EE2 is observed. This bioassay did not react to cumulative risk generated only by EDCs when cumulative risk of EH is low.
ERα-CALUX (option 2)	Compliance at 75% of sites. Non-compliance at 25% of sites	ERα-CALUX (option 2) could not forecast cumulative chemical risk of EH when it is low. This bioassay did not react to cumulative risk generated only by EDCs when cumulative risk of EH is low.
LiBERA	Compliance at 75% of sites. Non-compliance at 25% of sites	LiBERA showed a compliance with cumulative risk created by the EH only at sites having also an elevated cumulative risk due to EDCs. It could not identify cumulative risks posed only by EH at sites without or with a low cumulative risk from EDCs.

increased binding affinity for E2, EE2, BPA and 4-NP, making it a promising assay for detecting EDCs.

In the case of bioassays for which the EBT option 1 value was calculated based on averaged REP (INDIGO), it is strongly recommended the derivation of specific REP.

For effective WFD monitoring, the JRC recommends a combined assessment using EBM and chemical analysis as discussed in the literature (Neale et al., 2023b), considering the different sensitivity and specificity of the estrogenicity EBMs. EBMs should be used as an initial screening method, followed by chemical analysis to identify substances triggering the estrogenic response. Integrating EBM and chemical data would provide a comprehensive water quality assessment covering both known and unknown contaminants. Successful EBM implementation requires selecting appropriate assays - sensitive, reproducible, and relevant to all estrogenic contaminants occurring in the WS. Appropriate sampling (e.g., sampling locations, times, and frequencies) is essential to ensure that the WS are representative of the water body being monitored. Sample handling (e.g., containers, stabilisation, storage) and processing steps (e.g., filtration, extraction, and concentration steps to isolate the contaminants of interest) should be carried out under SOP. Quality controls – including positive/negative controls, replicates and standard reference materials – are critical for reliable results. EBM data need to be interpreted in the context of water quality standards and guidelines. This involves comparing the EBM results to EBT values that indicate potential risks to human health and the environment. Implementation should be iterative, incorporating feedback from Member States to refine the methods and address challenges.

To fully validate the EBT values proposed here, a second inter-laboratory study is planned as follow up. Ideally, involving all Member States and a larger number of freshwater samples from diverse EU locations, to ensure statistical robustness and account for climatic variability.

This study identified three areas of limitations: chemical analysis relied only on one laboratory (except for BPA); EEQ derivation methods varied across laboratories and in some cases lacked complete methodological detail, affecting comparability; and compliance comparability between EBMs and chemical analysis was tested only at eight sites with known exposure, a relatively simple context.

Future studies should address these limitations by engaging at least two laboratories for EH and EDCs concentration reporting within acceptable limits (e.g. 30% deviation); standardised EEQ derivation methods for consistency; and expanding comparability tests to more samples and diverse locations for robust verification.

5. Conclusions

This study provides scientific-based evidence for the application of EBMs and EBT within a legal framework and will contribute to the development of guidance for Member States to properly implement these methods in water quality monitoring under the WFD. Among these, estrogenicity EBMs are now ready for implementation, with some having been already identified as fit for purpose. They can be applied as initial screening monitoring prior to chemical analysis to identify the existence and cumulative (mixture) risk based on the MoA. In the case of estrogenicity monitoring, different EBMs could be used depending on the pollution source, since some of them proved to be more sensitive to estrogenic hormones, while others correlated better with other estrogenic EDCs.

The power of EBMs lies in its ability to capture the effects of chemical mixtures through their MoA. However, the coverage of MoA remains limited, and *in vitro* assays capable to detect multiple MoA – such as those critical for assessing per- and polyfluoroalkyl substances (PFAS) - are often lacking. Biotechnology-based innovations can play a key role in this field by accelerating efforts to address this gap.

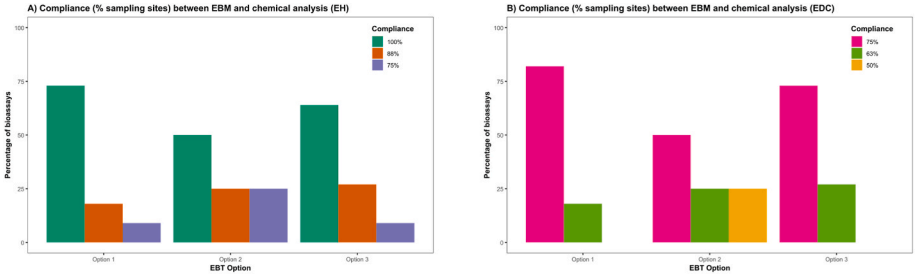

Fig. 6. The compliance percentages between EBM and chemical analyses of estrogenic hormones (EH) (A) and EDCs (B) for each EBT option across eight sampling sites (SP1-SP8) are presented. The results are shown as the percentage of bioassays achieving 75 %, 87.5% and 100 % compliance for all the SPs in the case EH or 50 %, 62.5 % and 75 % compliance in the case of estrogenic EDCs.

In conclusion, the implementation of the estrogenicity EBMs in the WFD will accelerate and foster biotechnology-based applications in the development of *in vitro* assays, thereby enhancing environmental monitoring and regulatory compliance.

CRediT authorship contribution statement

Livia Gómez: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Dimitar Marinov:** Writing – review & editing, Writing – original draft, Visualization, Formal analysis, Data curation. **Elena Porcel-Rodríguez:** Writing – review & editing, Methodology, Investigation. **Diletta Scaccabarozzi:** Writing – review & editing, Methodology, Investigation. **Isabella Sanseverino:** Writing – review & editing, Methodology, Investigation. **Pietro Sciuto:** Writing – review & editing, Methodology, Investigation. **Frank Narendja:** Writing – review & editing, Investigation, Formal analysis. **Romana Hornek-Gausterer:** Writing – review & editing, Investigation, Formal analysis. **Katharina Lenz:** Writing – review & editing, Investigation, Formal analysis. **Carole Chalon:** Writing – review & editing, Investigation, Formal analysis. **Yves Marneffe:** Writing – review & editing, Investigation, Formal analysis. **Přemysl Soldán:** Writing – review & editing, Investigation, Formal analysis. **Martin Hora:** Writing – review & editing, Investigation, Formal analysis. **Matti t. Leppänen:** Writing – review & editing, Investigation, Formal analysis. **Johanna Järvistö:** Writing – review & editing, Investigation, Formal analysis. **Jyrki Viidanoja:** Writing – review & editing, Investigation, Formal analysis. **Selim Ait-Aissa:** Writing – review & editing, Investigation, Formal analysis. **Emmanuelle Maillot-Marechal:** Writing – review & editing, Investigation, Formal analysis. **Abd El Rahman El Mais:** Writing – review & editing, Investigation, Formal analysis. **Beate I. Escher:** Writing – review & editing, Investigation, Formal analysis. **Maria König:** Writing – review & editing, Investigation, Formal analysis. **Henner Hollert:** Writing – review & editing, Investigation, Formal analysis. **Simone Wollenweber:** Writing – review & editing, Investigation, Formal analysis. **Sebastian Buchinger:** Writing – review & editing, Investigation, Formal analysis. **Marina Ohlig:** Writing – review & editing, Investigation, Formal analysis. **Roberta Pedrazzani:** Writing – review & editing, Investigation, Formal analysis. **Johan Lundqvist:** Writing – review & editing, Investigation, Formal analysis. **Agneta Oskarsson:** Writing – review & editing, Investigation, Formal analysis. **Etienne Vermeirssen:** Writing – review & editing, Investigation, Formal analysis. **Sibylle Maletz:** Writing – review & editing, Investigation, Formal analysis. **Daniel Olbrich:** Writing – review & editing, Investigation, Formal analysis. **Alvine C. Mehinto:** Writing – review & editing, Investigation, Formal analysis. **Kameron Wong:** Writing – review & editing, Investigation, Formal analysis. **Nancy Denslow:** Writing – review & editing, Investigation, Formal analysis. **Luca Amendola:** Writing – review & editing, Investigation, Formal analysis. **Ruben Gil-Solsona:** Writing – review & editing, Investigation, Formal analysis. **Pablo Gago-Ferrero:** Writing – review & editing, Investigation, Formal analysis. **Monica Potalivo:** Writing – review & editing, Investigation, Formal analysis. **Mario Carere:** Writing – review & editing, Investigation, Formal analysis. **Teresa Lettieri:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study has been funded by the Institutional Research program of the European Commission Joint Research Centre (JRC). The authors acknowledge Valeria Marchesi and Laura Tremolada from the Regional Environmental Protection Agency-Lombardy Region (ARPA-Lombardy) for their assistance in identifying appropriate sampling locations. The authors thank Roberta Carafa for her support during the sampling campaign. The authors are grateful to Simona Tavazzi, Giulio Mariani and Bernd Gawlik for their technical support on sample processing. We thank Niklas Wojtysiak, Andrea Schifferli, Nadine Bramaz and Jenny Braasch for experimental assistance. The authors would like to thank Eszter Simon (Swiss Federal Office for the Environment) for her participation and discussions in an advisory capacity and for revising the manuscript. The authors sincerely thank Luisa Colzani, Carola Forni, and Valeria Frattini (ARPA Lombardia – Dipartimento Regionale Prestazioni Analitiche) for their guidance and helpful suggestions on the chemical analysis of hormones. We gratefully acknowledge the invaluable support and dedication of our JRC colleagues, whose administrative and logistical contributions were essential to the successful completion of this study. BfG authors gratefully acknowledge financial support by the German Federal Ministry for the Environment, Climate Action, Nature Conservation and Nuclear Safety in line with the BMUNK-action 109. UFZ authors gratefully acknowledge access to the platform CIT-EPro (Chemicals in the Environment Profiler) funded by the Helmholtz Association for bioassay measurements and financial support from the Helmholtz POF IV Topic 9 “Healthy Planet-towards a non-toxic environment”.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envres.2026.124218.

Data availability

Data will be made available on request.

References

Adeel, M., Song, X., Wang, Y., Francis, D., Yang, Y., 2017. Environmental impact of estrogens on human, animal and plant life: a critical review. Environ. Int. 99, 107–119. https://doi.org/10.1016/j.envint.2016.12.010.

Altenburger, R., Brack, W., Burgess, R.M., Busch, W., Escher, B.I., Focks, A., et al., 2019. Future water quality monitoring: improving the balance between exposure and toxicity assessments of real-world pollutant mixtures. Environ. Sci. Eur. 31 (1), 1–17. https://doi.org/10.1186/s12302-019-0193-1.

Besselink, H., Brouwer, B., van der Burg, B., 2017. Validation and regulatory acceptance of bio-based approaches to assure feedstock, water & product quality in a bio-based economy. Ind. Crop. Prod. 106, 138–145. https://doi.org/10.1016/j.indcrop.2016.11.026.

Brion, F., De Gussem, V., Buchinger, S., Hollert, H., Carere, M., Porcher, J.M., et al., 2019. Monitoring estrogenic activities of waste and surface waters using a novel in vivo zebrafish embryonic (EASZY) assay: Comparison with in vitro cell-based assays and determination of effect-based trigger values. Environ. Int. 130, 104896. https://doi.org/10.1016/j.envint.2019.06.006.

Carvalho, R.N., Arukwe, A., Ait-Aissa, S., Bado-Nilles, A., Balzamo, S., Baun, A., et al., 2014. Mixtures of chemical pollutants at European legislation safety concentrations: how safe are they? Toxicol. Sci. 141 (1), 218–233. https://doi.org/10.1093/toxsci/kfu118.

Ciślak, M., Kruszelnicka, I., Zembrzuska, J., Ginter-Kramarczyk, D., 2023. Estrogen pollution of the European aquatic environment: a critical review. Water Res. 229, 119413. https://doi.org/10.1016/j.watres.2022.119413.

Drakvik, E., Altenburger, R., Aoki, Y., Backhaus, T., Bahadori, T., Barouki, R., et al., 2020. Statement on advancing the assessment of chemical mixtures and their risks for human health and the environment. Environ. Int. 134, 105267. https://doi.org/10.1016/j.envint.2019.105267.

EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02000L0060-20141120.

EC, 2008. Directive 2008/105/EC of the European Parliament and of the council of 16 December 2008 on environmental quality standards in the field of water policy. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02008L0105-20130913&qid=1764256554981.

EC, 2022. European commission, proposal directive October 2022. Proposal for a directive of the European parliament and of the council. Available on-line here: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0540.

Escher, B.I., Aït-Aïssa, S., Behnisch, P.A., Brack, W., Brion, F., Brouwer, A., et al., 2018. Effect-based trigger values for in vitro and in vivo bioassays performed on surface water extracts supporting the environmental quality standards (EQS) of the european Water Framework Directive. Sci. Total Environ. 628, 748–765. https://doi.org/10.1016/j.scitotenv.2018.01.340.

Escher, B.I., Neale, P.A., 2021. Effect-Based trigger values for mixtures of chemicals in surface water detected with In vitro bioassays. Environ. Toxicol. Chem. 40, 487–499. https://doi.org/10.1002/etc.4944.

EU, 2013. Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0039.

EU, 2015. Commission Implementing Decision (EU) 2015/495 of 20 March 2015 establishing a watch list of substances for union-wide monitoring in the field of water policy. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015D0495.

EU, 2018. Commission Implementing Decision (EU) 2018/840 of 5 June 2018 establishing a watch list of substances for union-wide monitoring in the field of water policy. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018D0840.

Ferrero, V.E.V., Pedotti, M., Chiadò, A., Simonelli, L., Calzolai, L., Varani, L., Lettieri, T., 2014. Rational modification of estrogen receptor by combination of computational and experimental analysis. PLoS One 9 (7), e102658. https://doi.org/10.1371/journal.pone.0102658.

Glineur, A., Nott, K., Carbonnelle, P., Ronkart, S., Pollet, T., Purcaro, G., 2024. Occurrence and environmental risk assessment of 4 estrogenic compounds in surface water in Belgium in the frame of the EU Watch list. Environ. Sci. Pollut. Control Ser. 31 (5), 6857–6873. https://doi.org/10.1007/s11356-023-31698-9.

Gómez, L., Niegowska, M., Navarro, A., Amendola, L., Arukwe, A., Ait-Aissa, S., et al., 2021. Estrogenicity of chemical mixtures revealed by a panel of bioassays. Sci. Total Environ. 785, 147284. https://doi.org/10.1016/j.scitotenv.2021.147284.

Gonsioroski, A., Mourikes, V.E., Flaws, J.A., 2020. Endocrine disruptors in water and their effects on the reproductive system. Int. J. Mol. Sci. 21 (6), 1929. https://doi.org/10.3390/ijms21061929. PMID: 32178293; PMCID: PMC7139484.

ISO Standard, 2018a. ISO 19040-1:2018, Water Quality — Determination of the Estrogenic Potential of Water and Waste Water Part 1: Yeast Estrogen Screen (Saccharomyces cerevisiae).

ISO Standard, 2018b. ISO 19040-2:2018, Water Quality — Determination of the Estrogenic Potential of Water and Waste Water. Part 2: Yeast estrogen screen (A-YES, Arxula adeninivorans).

ISO Standard, 2018c. ISO 19040-3:2018, Water Quality — Determination of the Estrogenic Potential of Water and Waste Water. Part 3: in Vitro Human Cell-based Reporter Gene Assay.

Jarošová, B., Bláha, L., Giesy, J.P., Hilscherová, K., 2014a. What level of estrogenic activity determined by in vitro assays in municipal waste waters can be considered as safe? Environ. Int. 64, 98–109. https://doi.org/10.1016/j.envint.2013.12.009.

Jarošová, B., Erseková, A., Hilscherová, K., Loos, R., Gawlik, B.M., Giesy, J.P., Bláha, L., 2014b. Europe-wide survey of estrogenicity in wastewater treatment plant effluents: the need for the effect-based monitoring. Environ. Sci. Pollut. Res. 21, 10970–10982. https://doi.org/10.1007/s11356-014-3056-8.

Könemann, S., Kase, R., Simon, E., Swart, K., Buchinger, S., Schlüsener, M., et al., 2018. Effect-based and chemical analytical methods to monitor estrogens under the European water framework directive. TrAC, Trends Anal. Chem. 102, 225–235. https://doi.org/10.1016/j.trac.2018.02.008.

Kunz, P.Y., Kienle, C., Carere, M., Homazava, N., Kase, R., 2015. In vitro bioassays to screen for endocrine active pharmaceuticals in surface and waste waters. J. Pharmaceut. Biomed. Anal. 106, 107–115. https://doi.org/10.1016/j.jpba.2014.11.018.

Kunz, P.Y., Simon, E., Creusot, N., Jayasinghe, B.S., Kienle, C., Maletz, S., Vermeirssen, E. L., 2017. Effect-based tools for monitoring estrogenic mixtures: evaluation of five in vitro bioassays. Water Res. 110, 378–388. https://doi.org/10.1016/j.watres.2016.10.062.

Leusch, F.D., De Jager, C., Levi, Y., Lim, R., Puijker, L., Sacher, F., et al., 2010. Comparison of five in vitro bioassays to measure estrogenic activity in environmental waters. Environ. Sci. Technol. 44 (10), 3853–3860. https://doi.org/10.1021/es903899d.

Leusch, F.D., Khan, S.J., Gagnon, M.M., Quayle, P., Trinh, T., Coleman, H., et al., 2014. Assessment of wastewater and recycled water quality: a comparison of lines of evidence from in vitro, in vivo and chemical analyses. Water Res. 50, 420–431. https://doi.org/10.1016/j.watres.2013.10.056.

Loos, R., Marinov, D., Sanseverino, I., Napierska, D., Lettieri, T., 2018. Review of the 1st Watch List Under the Water Framework Directive and Recommendations for the 2nd Watch List. Publications Office of the European Union, Luxembourg. https://doi.org/10.2760/614367.

Luijten, M., Vlaanderen, J., Kortenkamp, A., Antignac, J.P., Barouki, R., Bil, W., et al., 2023. Mixture risk assessment and human biomonitoring: lessons learnt from HBM4EU. Int. J. Hyg Environ. Health 249, 114135. https://doi.org/10.1016/j.ijheh.2023.114135.

Murk, A.J., et al., 2002. Detection of estrogenic potency in wastewater and surface water with three in vitro bioassays. Environ. Toxicol. Chem. 21 (1), 16–23. https://doi.org/10.1002/etc.5620210103.

Moermond, C., Venhuis, B., van Elk, M., Oostlander, A., van Vlaardingen, P., Marinkovic, M., van Dijk, J., 2018. Cytostatics in Dutch Surface Water: Use, Presence and Risks to the Aquatic Environment. https://www.rivm.nl/bibliotheek/rapporten/2018-0067.pdf.

Neale, P.A., Brack, W., Aït-Aïssa, S., Busch, W., Hollender, J., Krauss, M., et al., 2018. Solid-phase extraction as sample preparation of water samples for cell-based and other in vitro bioassays. Environ. Sci. Process. Impacts 20 (3), 493–504. https://doi.org/10.1039/C7EM00555E.

Neale, P.A., Escher, B.I., de Baat, M.L., Enault, J., Leusch, F.D.L., 2023a. Effect-Based trigger values are essential for the uptake of effect-based methods in water safety planning. Environ. Toxicol. Chem. 42 (3), 714–726. https://doi.org/10.1002/etc.5544.

Neale, P.A., Escher, B.I., de Baat, M.L., Dechesne, M., Dingemans, M.M.L., Enault, J., Pronk, G.J., Smeets, P.W.M.H., Leusch, F.D.L., 2023b. Application of effect-based methods to water quality monitoring: answering frequently asked questions by water quality managers, regulators, and Policy makers. Environ. Sci. Technol. 57 (15), 6023–6032. https://doi.org/10.1021/acs.est.2c06365, 2023.

OECD, 2021. Performance-Based Test Guideline for Stably Transfected Transactivation in Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists, Test Guideline No. 455. Organisation for Economic Cooperation and Development, Paris.

Rutishauser, B.V., Pesonen, M., Escher, B.I., Ackermann, G.E., Aerni, H.R., Suter, M.J.F., Eggen, R.I., 2004. Comparative analysis of estrogenic activity in sewage treatment plant effluents involving three in vitro assays and chemical analysis of steroids. Environ. Toxicol. Chem. 23 (4), 857–864. https://doi.org/10.1897/03-286.

Savitz, D.A., Hattersley, A.M., 2023. Evaluating chemical mixtures in epidemiological studies to inform regulatory decisions. Environ. Health Perspect. 131 (4), 045001. https://doi.org/10.1289/EHP1189.

Schultis, T., Metzger, J.W., 2004. Determination of estrogenic activity by LYES-assay (yeast estrogen screen-assay assisted by enzymatic digestion with lyticase). Chemosphere 57 (11), 1649–1655. https://doi.org/10.1016/j.chemosphere.2004.06.027.

Schulze, T., Neale, P.A., Ahlheim, J., Beckers, L.M., König, M., Krüger, J., et al., 2024. A guidance for the enrichment of micropollutants from wastewater by solid-phase extraction before bioanalytical assessment. Environ. Sci. Eur. 36 (1), 165. https://doi.org/10.1186/s12302-024-00990-x.

Serra, H., Brion, F., Chardon, C., Budzinski, H., Schulze, T., Brack, W., Aït-Aïssa, S., 2020. Estrogenic activity of surface waters using zebrafish-and human-based in vitro assays: The Danube as a case-study. Environ. Toxicol. Pharmacol. 78, 103401. https://doi.org/10.1016/j.etap.2020.103401.

Simon, E., Duffek, A., Stahl, C., Frey, M., Scheurer, M., Tuerk, J., et al., 2022a. Biological effect and chemical monitoring of Watch List substances in European surface waters: steroidal estrogens and diclofenac–effect-based methods for monitoring frameworks. Environ. Int. 159, 107033. https://doi.org/10.1016/j.envint.2021.107033.

Simon, E., Riegraf, C., Schifferli, A., Olbrich, D., Bucher, T., Vermeirssen, E.L.M., 2022b. Evaluation of three ISO estrogen receptor transactivation assays applied to 52 domestic effluent samples. Environ. Toxicol. Chem. 41, 2512–2526. https://doi.org/10.1002/etc.5445.

Singh, P.K., Kumar, U., Kumar, I., Dwivedi, A., Singh, P., Mishra, S., et al., 2024. Critical review on toxic contaminants in surface water ecosystem: sources, monitoring, and its impact on human health. Environ. Sci. Pollut. Control Ser. 31 (45), 56428–56462. https://doi.org/10.1007/s11356-024-34932-0.

Tang, J.Y., Escher, B.I., 2014. Realistic environmental mixtures of micropollutants in surface, drinking, and recycled water: herbicides dominate the mixture toxicity toward algae. Environ. Toxicol. Chem. 33 (6), 1427–1436. https://doi.org/10.1002/etc.2580.

van der Oost, R., Sileno, G., Suárez-Muñoz, M., Nguyen, M.T., Besselink, H., Brouwer, A., 2017. SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality assessment: part I–model design and effect-based trigger values. Environ. Toxicol. Chem. 36 (9), 2385–2399. https://doi.org/10.1002/etc.3836.

Wagner, M., Vermeirssen, E.L., Buchinger, S., Behr, M., Magdeburg, A., Oehlmann, J., 2013. Deriving bio-equivalents from in vitro bioassays: assessment of existing uncertainties and strategies to improve accuracy and reporting. Environ. Toxicol. Chem. 32 (8), 1906–1917. https://doi.org/10.1002/etc.2256.

Wernersson, A.S., Carere, M., Maggi, C., Tusil, P., Soldan, P., James, A., et al., 2015. The European technical report on aquatic effect-based monitoring tools under the water framework directive. Environ. Sci. Eur. 27 (1), 7. https://doi.org/10.1186/s12302-015-0039-4.

Zhao, J.L., Ying, G.G., Chen, F., Liu, Y.S., Wang, L., Yang, B., et al., 2011. Estrogenic activity profiles and risks in surface waters and sediments of the Pearl River system in South China assessed by chemical analysis and in vitro bioassay. J. Environ. Monit. 13 (4), 813–821. https://doi.org/10.1039/C0EM00473A.