Biological effect and chemical monitoring of Watch List substances in European surface waters: Steroidal estrogens and diclofenac – Effect-based methods for monitoring frameworks

Eszter Simon, Anja Duffek, Cordula Stahl, Manfred Frey, Marco Scheurer, Jochen Tuerk, Linda Gehrmann, Sarah Könemann, Kees Swart, Peter Behnisch, et al.

Environment International 159 (2022) 107033

Contents lists available at ScienceDirect

Environment International

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

Biological effect and chemical monitoring of Watch List substances in European surface waters: Steroidal estrogens and diclofenac – Effect-based methods for monitoring frameworks

Eszter Simon [a,*,1], Anja Duffek [b], Cordula Stahl [c], Manfred Frey [c], Marco Scheurer [d], Jochen Tuerk [e], Linda Gehrmann [e], Sarah Könemann [f], Kees Swart [g], Peter Behnisch [h], Daniel Olbrich [a], François Brion [i], Selim Aït-Aïssa [i], Robert Pasanen-Kase [i], Inge Werner [a], Etiënne L.M. Vermeirssen [a,*]

[a] Swiss Centre for Applied Ecotoxicology, Dübendorf, Switzerland
[b] German Environment Agency (UBA), Berlin, Germany
[c] Steinbeis-Innovationszentrum Zellkulturtechnik, c/o University of Applied Sciences Mannheim, Germany
[d] TZW: DVGW-Technologiezentrum Wasser (German Water Centre), Karlsruhe, Germany
[e] Institut für Energie- und Umwelttechnik e. V. (IUTA, Institute of Energy and Environmental Technology), Duisburg, Germany
[f] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland
[g] BioDetection Systems B.V., Amsterdam, the Netherlands
[h] National Institute of Industrial Environment and Risks (INERIS), UMR-I 02 SEBIO, Verneuil-en-Halatte, France
[i] State Secretariat for Economic Affairs (SECO), Labour Directorate Section Chemicals and Work, Bern, Switzerland

ARTICLE INFO

Handling Editor: Frederic Coulon

Keywords:
Water Framework Directive (WFD)
ERα-CALUX
A-YES
EASZY
COX-inhibition
Non-steroidal anti-inflammatory drug (NSAID)
Standard addition
Predicted no-effect concentration (PNEC)

ABSTRACT: RATIONALE OF THE STUDY

Three steroidal estrogens, 17α-ethinylestradiol (EE2), 17β-estradiol (E2), estrone (E1), and the non-steroidal anti-inflammatory drug (NSAID), diclofenac have been included in the first Watch List of the Water Framework Directive (WFD, EU Directive 2000/60/EC, EU Implementing Decision 2015/495). This triggered the need for more EU-wide surface water monitoring data on these micropollutants, before they can be considered for inclusion in the list of priority substances regularly monitored in aquatic ecosystems. The revision of the priority substance list of the WFD offers the opportunity to incorporate more holistic bioanalytical approaches, such as effect-based monitoring, alongside single substance chemical monitoring. Effect-based methods (EBMs) are able to measure total biological activities (e.g., estrogenic activity or cyxlooxygenase [COX]-inhibition) of specific group of substances (such as estrogens and NSAIDs) in the aquatic environment at low concentrations (pg/L). This makes them potential tools for a cost-effective and ecotoxicologically comprehensive water quality assessment. In parallel, the use of such methods could build a bridge from chemical status assessments towards ecological status assessments by adressing mixture effects for relevant modes of action. Our study aimed to assess the suitability of implementing EBMs in the WFD, by conducting a large-scale sampling and analysis campaign of more than 70 surface waters across Europe. This resulted in the generation of high-quality chemical and effect-based monitoring data for the selected Watch List substances. Overall, water samples contained low estrogenicity (0.01–1.3 ng E2-Equivalent/L) and a range of COX-inhibition activity similar to previously reported levels (12–1600 ng Diclofenac-Equivalent/L). Comparison between effect-based and conventional analytical chemical methods showed that the chemical analytical approach for steroidal estrogens resulted in more (76%) non-quantifiable data, i.e., concentrations were below detection limits, compared to the EBMs (28%). These results demonstrate the excellent and sensitive screening capability of EBMs.

* Corresponding authors.
E-mail addresses: eszter.simon@bafu.admin.ch (E. Simon), etienne.vermeirssen@oekotoxzentrum.ch (E.L.M. Vermeirssen).
[1] Present address: Federal Office for the Environment FOEN, Air Pollution and Chemicals Division, Industrial Chemicals Section, Bern, Switzerland.

https://doi.org/10.1016/j.envint.2021.107033
Received 23 September 2021; Received in revised form 2 December 2021; Accepted 3 December 2021
Available online 31 December 2021

1. Introduction: Effect-based water quality assessment

1.1. Watch List mechanism under the EU Water Framework Directive

The EU Water Framework Directive (WFD) provides a regulatory framework for ecological and chemical status assessment of EU surface waters. The WFD covers 45 prioritized chemicals that member states are required to regularly monitor using chemical analysis (EU, Directive, 2000; EU, Directive, 2013; EU, Commission Implementing Decision (EU), 2015; EU, Commission Implementing Decision (EU), 2018). In recent years, it has often been questioned if this list is up to date and relevant. Concerns were raised about the increasing amount and variety of chemicals entering and affecting the aquatic ecosystem. Many of these chemicals are still widely neglected in monitoring (Altenburger et al., 2015; Brack, 2015; Carere et al., 2015). To support the identification of new priority substances, a Watch List mechanism was introduced under the WFD in 2013 (EU, Directive, 2013). The ultimate goal was to generate sufficient high-quality monitoring data on potential water pollutants that may pose a risk to or via the aquatic environment. Steroidal estrogens and diclofenac, among others, were listed on the first Watch List published in 2015 for a continuous monitoring for up to four years (EU, Commission Implementing Decision (EU), 2015). This Watch List is submitted to regular review and revision.

1.2. Incorporating effect-based screening methods (EBMs) into water quality assessment frameworks

Besides identifying potential and relevant chemicals, a revision of the WFD's priority list provides an opportunity to recognize and accommodate more comprehensive monitoring approaches and acknowledge a "paradigm shift" towards effect-based water quality assessment. Numerous studies (Leusch et al., 2010; Hamers et al., 2013; Altenburger et al., 2015; Brack et al., 2016; Busch et al., 2016; Brack et al., 2017; van der Oost et al., 2017b; Kase et al., 2018; Könemann et al., 2018; De Baat et al., 2019) have drawn attention to the effectiveness of effect-based screening methods (bioassays) and thereby their suitability in monitoring frameworks. Bioassays are already used in pharmaceutical screening, food safety assessment and chemical regulation but have only sporadically entered water regulation, particularly in Europe (Combes et al., 2004; EU, Regulation (EU) No 589/, 2014; Severin et al., 2017).

Bioassays offer several advantages over chemical target monitoring and incorporating them in monitoring programs is relatively straightforward. First, chemical analytical monitoring surveys routinely measure concentrations of single target compounds. However, observed toxic effects in surface waters can often not or only partially be ascribed to these regularly measured chemicals (Busch et al., 2016; Neale et al., 2017a, 2017b; De Baat et al., 2019). To decide, whether known and unknown emerging chemicals pose risks to organisms in aquatic environments, one has to be aware of their mixture effects, metabolites, and transformation products with the same mode of action (MOA). Complex mixtures of compounds with the same MOA could lead to physiological responses that deviate quantitatively from the additive physiological responses of the individually measured compounds (Bernhard et al., 2017). Second, certain chemicals (such as steroidal estrogens) are bioactive at very low concentrations (pg/L), challenging the detection capability of the existing routine analytical methods (Kunz et al., 2015; Kase et al., 2018; Könemann et al., 2018; Loos et al., 2018b). EBMs can overcome these limitations by revealing the combined effect of targeted, not-targeted chemicals and also their mixture with satisfactory detection limits (i.e., lower than the respective limit values set). Thus, bioassays offer support for a comprehensive evaluation of environmental status. Third, besides revealing the actual biological effect of chemical mixtures, they can also serve as a cost-effective prescreening tool and early-warning system. After the discovery of ecotoxicological effect-driven hotspots, further chemical and pollution source identification can be conducted resulting in a comprehensive impact assessment strategy. Lastly, and most importantly, a strategy involving prescreening with highly sensitive bioassays could prevent costly and complex chemical analyses at sites with low ecotoxicological risks (Connon et al., 2012; van der Oost et al., 2017a; De Baat et al., 2019).

1.3. The project – Background and European collaborative effort

Between 2015 and 2016, a first study was conducted, in which a screening of steroidal estrogens was performed on European wastewater treatment plant (WWTP) effluents and their receiving surface waters. The collected samples were analyzed by different laboratories across Europe to compare five EBMs with three LC-MS/MS methods (Kase et al., 2018; Könemann et al., 2018; Brion et al., 2019). The water samples were predominantly collected from contaminated sites to enable the methods comparison at elevated (i.e., robustly detectable) concentrations, thus avoiding biases associated with different method-specific LODs. Significant estrogenicity was detected in most of the samples and the results were comparable between the methods. The study revealed the applicability of EBMs as prescreening tool to identify hotspots (i.e., sites with exceedances of thresholds) and as valuable complements to chemical analysis.

The work described here aims to expand on these findings and assess the practical feasibility of an effect-based approach on a larger scale. We increased the number of sampling locations, targeted an additional set of chemicals (diclofenac and other non-steroidal anti-inflammatory drugs, NSAIDs), and reduced the number of measuring techniques applied. Our sampling campaign included 73 surface water monitoring stations, located in 14 different European countries. This ensured a broader and a more neutral selection of representative European surface water matrices, not only targeting hotspots. Our approach allowed us to test effect-based methods for two groups of chemicals of environmental concern. Steroidal estrogens are particularly harmful to aquatic organisms even at trace concentrations, altering their hormonal system and ultimately leading to sexual developmental and behavioral effects (Jobling et al., 2006; Kidd et al., 2007; Söffker and Tyler, 2012; Adeel et al., 2017). It is well known that estrogens enter surface waters via WWTP effluents and from agricultural areas, and therefore have already been widely investigated (Gaulke et al., 2009; Itzel et al., 2017; Gehrmann et al., 2018). Similarly, diclofenac and other NSAIDs frequently occur in surface waters, being ineffectively eliminated by WWTPs, but also as discharge from intensive farming in rural areas (Sathishkumar et al., 2020). Due to the phylogenetic conservation of central signal transduction pathways among vertebrates, NSAIDs affect vertebrate species, fish in particular, even at low concentrations (Bernhard et al., 2017). For both, steroidal estrogens and NSAIDs, *in vitro* bioassays are readily available thus they constituted excellent candidates for this study.

In contrast to our first study, fewer techniques were employed. LC-MS/MS with internal standard calibration remained as reference method, against which results of EBMs were assessed. For estrogenic activity, two ISO standardized and commercially available *in vitro* assays were chosen: a yeast-based estrogen screen (A-YES, (Hettwer et al., 2018; ISO 19040-2:2018, Water quality) and a human cell line-based estrogen receptor transactivation assay (ERα-CALUX, (Sonneveld et al., 2005; van der Linden et al., 2008; Simon et al., 2019; ISO 19040-3:2018, Water quality). Both these methods have been extensively validated with environmental waters (Gehrmann et al., 2018; ISO 19040-1:2018, Water quality). The ERα-CALUX was one of the EBMs investigated in the first study (Könemann et al., 2018). Both assays measure the gene-induced production of a specific enzyme due to a transcriptional cascade triggered by estrogenic chemicals upon their binding to the human estrogen receptor (engineered in the cells). This reporter enzyme transforms the substrate present in the assay medium and produces a quantifiable luminescence or fluorescence response in a proportional manner (Sonneveld et al., 2005; Hettwer et al., 2018). A

selected set of samples were also analyzed with the *in vivo* zebrafish embryonic assay (EASZY, (Brion et al., 2019), to establish a link between *in vitro* and possible adverse effects *in vivo* (at the organism level). Such a link supports the ecotoxicological relevance of cell-based assays and their predictive capacity. Cyclooxygenase (COX)-inhibition was tested with the rapid mammalian biosensor cell line-based *in vitro* assay that delivers information on pharmaceutical - target molecule interaction within minutes. NSAIDs exert their analgesic action through obstruction of special COX enzyme(s) that are involved in prostaglandin synthesis (the chemical promoting inflammation, pain and fever) (Gan, 2010). In this assay, cyclooxygenase-1 functions as sensor for COX-inhibition and a redox sensitive green fluorescent protein as reporter of COX-activity. The co-expression of sensor and reporter proteins in the sensor cell line allows the monitoring of COX-activity in real-time (Bernhard et al., 2017). The COX-inhibition (or NSAID) biosensor has been proven to be robust, reliable and suitable for environmental water testing (e.g., Bernhard et al., 2017; Frey and Scheurer, 2020).

2. Material and methods

2.1. Collaborative research endeavor to collect and analyze surface water samples

Between autumn 2017 and spring 2018, operators from the participating European countries (13 EU Member States: AT, BE, CZ, DE, DK, FI, FR, LV, NL, SCO, SE, SK, SL and four cantons from Switzerland) collected surface water samples following a standardized sampling and transport protocol, but without predefined criteria for the selection of sampling locations. The majority of sampling stations were routine WFD monitoring stations. As supporting information on the characteristics of sampling locations, project partners estimated the local impact by agricultural runoff and discharge from wastewater treatment plants (WWTP) as low, moderate, or high. However, to preserve anonymity, no systematic information (such as size and geographic proximity of WWTP effluent discharges and industrial or agricultural areas) was provided for most of the sampling sites (Table S1 in Supplementary data).

The collaborative study design is depicted in Scheme 1. A total of 75 surface water samples (~5.5 L each) were shipped directly to the central laboratory (NL), where the preparation and distribution of the samples were performed. Ten field blanks (ultra-pure water) were also added to the sample series of six randomly selected countries, and were handled the same way. Encoded aliquots of the sample extracts were transported to project partner laboratories (CH, NL, DE, FR) having the facilities to analyze steroidal estrogens and diclofenac along with other pharmaceuticals (COX-inhibitors, e.g., ibuprofen).

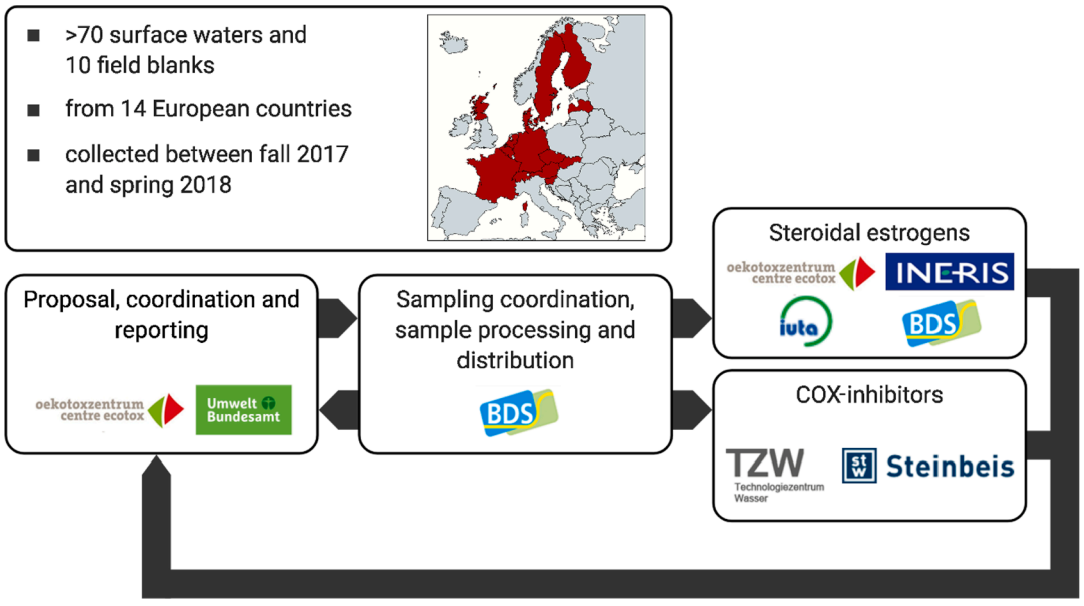
2.2. Sample preparation, bioassay, and LC-MS/MS analyses

2.2.1. Steroidal estrogens

Sample preparation and ERα-CALUX were conducted in The Netherlands (NL), A-YES in Germany (DE) and LC-MS/MS in Switzerland (CH). The measurements, data evaluation and reporting were executed according to the routine practice of the analyzing laboratory.

Water samples (4 L) were collected in pre-cleaned (three times methanol rinsed) 5.5 L aluminium bottles, pH was adjusted to 3 with HCl, and kept frozen at −20 °C until used. Prior to extraction, each sample was thawed overnight at room temperature, filtered (2.7 μm glass fibre filters, Milipore), solid-phase extracted (SPE) using C18 cartridges and silica SPE clean-up according to the protocol described by Könemann et al. (2018) with a few modifications. This time, each water sample was divided in 1 L aliquots, each extracted on a single C18 cartridge (i.e., four cartridges per sample, Phenomenex Strata C18-E, 500 mg), resulting in 4 mL extracts with an enrichment factor of 1000. Clean-up was performed on Silica (CHROMABOND SiOH, Macherey-Nagel, 1000 mg) previously heat-activated columns (one column per mL extract). Cleaned ethanolic extracts were pooled together, then split into four 1 mL aliquots and sent to project partners for analysis. This ensured that different analyses were performed with the same sample extracts.

Quantification of 17α-ethinylestradiol (EE2), 17β-estradiol (E2), estrone (E1) was done using an Agilent G6495A Triple Quadrupole (QQQ) mass spectrometer coupled with an UHPLC, after external calibration with matching deuterium-labelled analytes. The method is described in detail in our preceding studies (Könemann et al., 2018; Simon et al., 2019). The LOQ for each steroidal estrogen was 0.1 ng/L implying three times lower LOD of 0.03 ng/L. Measured individual steroidal estrogen concentrations in the samples were compared to the



Scheme 1. Study design for monitoring selected Watch List substances: steroidal estrogens (E1, E2, EE2) and COX-inhibitors (e.g., diclofenac) with centralized sample workup and distribution. The logos represent the responsible institutes. Map chart was created and customized on https://mapchart.net/ and sampled countries are marked red. The figure was created with BioRender.com.

respective predicted no-effect concentrations (PNEC) or maximal acceptable method detection limits (MDL) to identify exceedances: E1 (3.6 ng/L), E2 (0.4 ng/L) and EE2 (0.035 ng/L) (EU, Commission Implementing Decision (EU), 2015; Loos et al., 2018b).

Additionally, half of the surface water extract aliquots (sent to and analyzed in CH) were spiked with a mixture of EE2, E2 and E1 at their LOQ concentration level of 0.1 ng/L. This standard addition - visualized in Figure S1 - served primary as a quality control of the LOQ concentration levels of the target analytes. Besides that, it aimed to better characterize concentrations below LOQ and eventually enable a more accurate ecotoxicological risk assessment and a robust data evaluation (i.e., prediction of combined effects based on individual estrogen composition of the sample).

Effect concentrations were determined in two standardized *in vitro* estrogen receptor transactivation assays (ERTAs): the human cell-based ERα-CALUX and the yeast estrogen screen A-YES (ISO 19040-2:2018, Water quality; ISO 19040-3:2018, Water quality). Estrogenic activity of the extracts was reported by the respective labs (NL, DE) and expressed as reference compound, E2-equivalent concentration (ng EEQ/L water sample). EEQs were compared to the proposed effect-based trigger value (EBT) of 0.4 ng/L based on the PNEC of the reference, E2 to identify samples with exceedances (Loos, 2012; Kunz et al., 2015; Wernersson et al., 2015; Kase et al., 2018). Samples were tested in ERα-CALUX at a maximum relative enrichment of ~ 20. The sample-specific quantification limits (LOQ) in ERα-CALUX were reported by the performing laboratory. It was defined based on the effect level of the sample used for deriving EEQ concentration and adjusted for the highest enrichment factor tested. LOQ in the ERα-CALUX ranged from 0.01 to 0.05, on average 0.02 ng EEQ/L. In A-YES, detection limits (LODs) were defined and reported based on the method calibration and solvent control measurements (i.e. not sample-specific). Samples were analyzed at a maximum relative enrichment of 500. LOD in A-YES ranged from 0.03 to 0.05, on average 0.04 ng EEQ/L.

To confirm the toxicological relevance of *in vitro*-based estrogenicity assessments, a smaller set of samples was run in the *in vivo* EASZY assay. Twenty extracts (1 mL ethanolic aliquots) together with one of the ten field blanks were sent for *in vivo* estrogenicity analysis (FR) with zebrafish embryos (EASZY assay, (Brion et al., 2019). The samples were concentrated and resuspended in 0.1 mL or 0.05 mL of dimethylsulfoxide (DMSO) which was further diluted 1:1000 in exposure medium. Samples were analysed at a maximum relative enrichment factor (REF) up to 20. Effect concentrations in ng EEQ/L as well as LODs (3.2 ng EEQ/L) were reported by the analyzing laboratory.

2.2.2. Diclofenac and other NSAIDs

Both sample preparation and (effect-based and chemical) analyses, as well as data evaluation and reporting were conducted in Germany (DE) following routine protocols of the analyzing laboratories.

Pre-cleaned (three times methanol rinsed) 1.25 L aluminum bottles were filled with 1 L water samples without acidification (39 of the 71 surface waters and five of the ten field blanks) and stored at 4 °C until used. Samples were adjusted to pH 3 with hydrochloric acid and divided into sub-samples of 700 mL and 250 mL. Sub-samples for chemical analysis (250 mL) were spiked with appropriate isotope standards. Fortification with internal standards was avoided for sub-samples (700 mL) prepared for application in the bioassay. All sub-samples were pre-concentrated with a polymeric sorbent material (Strata-X, 200 mg, Phenomenex). Prior to enrichment, SPE cartridges were washed with 2x3 mL methanol and 2x3 mL distilled water (pH 3). After SPE, the sorbent was dried under a gentle stream of nitrogen and eluted with 2x2.5 mL methanol and 2.5 mL acetone. Both extracts were evaporated till dryness under a gentle stream of nitrogen.

Quantification of diclofenac and other NSAIDs was done by internal standard calibration using a model 1200 SL HPLC system from Agilent Technologies coupled to an API 4000 Q-Trap triple-quadrupole mass spectrometer according to the method described by Bernhard *et al.* (Bernhard et al., 2017). The dry residue of the 250 mL sub-sample was reconstituted in 200 µL for the chemical analysis resulting in an enrichment factor of 1250. Average LOQ for diclofenac and the other NSAIDs was 12 ng/L. Measured diclofenac concentrations were compared to its PNEC of 50 ng/L (Loos, 2012; EU, Commission Implementing Decision (EU), 2015).

Effect concentrations were determined in the *in vitro* cell-based NSAID assay (DE) according to the method described by Bernhard *et al.* (Bernhard et al., 2017) and in the UBA (German Environment Agency) report of the project "Validation of the NSAID *in vitro* assay for biomonitoring of NSAID activities in surface waters" (Frey and Scheurer, 2020). For the NSAID assay, the dry residue of the 700 mL sub-sample was reconstituted in a HEPES buffer. The biosensor cell line expressed the genetically encoded fluorescent redox sensitive green fluorescent protein (GFP) together with human COX-1.

COX-inhibition was expressed as reference compound, diclofenac-equivalent concentration (DicEQ, ng/L water sample). DicEQs were then compared to the suggested diclofenac PNEC of 50 ng/L (Loos et al., 2018b). The LOD of the NSAID *in vitro* assay (i.e., referring to a 150times relative enrichment) was 6 ng/L.

2.3. Comparison between LC-MS/MS and EBMs

Individual chemical concentrations of three estrogens (i.e., E1, E2, EE2) were multiplied by their bioassay-specific relative potency factors (REP; i.e., the potency relative to the reference compound, E2) and summed to obtain the total nominal "chemical measurement-based", calculated biological effect concentrations for each sample, EEQ_{chem} (Eq. (1)) following the concept described by Kunz *et al.* (Kunz et al., 2017).

This allowed direct comparison of LC-MS/MS (i.e., calculated $EEQ_{chem\ A\text{-}YES}$ or $EEQ_{chem\ CALUX}$) to each measured biological effect concentrations ($EEQ_{bio\ A\text{-}YES}$ or $EEQ_{bio\ CALUX}$) (Fig. 1.)

$$EEQ_{chem\ A-YES\ or\ CALUX} = \sum \text{Conc. } (ng/L)_{E1\ or\ E2\ or\ EE2} \times REP_{in\ the\ A-YES\ or\ CALUX\ of\ E1\ or\ E2\ or\ EE2} \quad (1)$$

Values below LOD were zeroed for the EEQ_{chem} calculation. Not robustly quantifiable (above LOD, but below LOQ) LC-MS/MS concentrations were still considered and used for EEQ_{chem} calculations.

In the "Supplementary data", the estrogenic relative potencies for the estrogenic bioassays are summarized (Table S2). As mainly diclofenac and no other NSAIDs were found in the samples (see Results and discussion), *in vitro* effect concentrations could be directly compared to the diclofenac concentrations (without considering the relative potency of other NSAIDs).

2.4. Method validity

Field blanks (i.e., procedure blank) consisted of ultra-pure water that underwent each step of the analyses, including transport, extraction, clean-up and analyses. Four samples and the field blank from one of the participating countries were excluded from data evaluation, due to an E2 contamination detected both by LC-MS/MS and ERTAs. These samples were removed from the data set and 71 samples remained. Neither target analytes nor estrogenicity or COX-inhibition were detected in the remaining field blanks.

Positive control samples consisted of ultra-pure water spiked with the respective target analytes at different concentration levels (E2 for estrogen monitoring and diclofenac for NSAID monitoring) to determine the accuracy of the effect-based methods. This was crucial as the application of internal standards in bioassays is not recommended, since they may contribute to the observed responses and lead to overestimation of the total biological activities (Simon et al., 2015).

These positive controls were not intended to verify analyte stability over the entire sample handling process. The stability of compounds

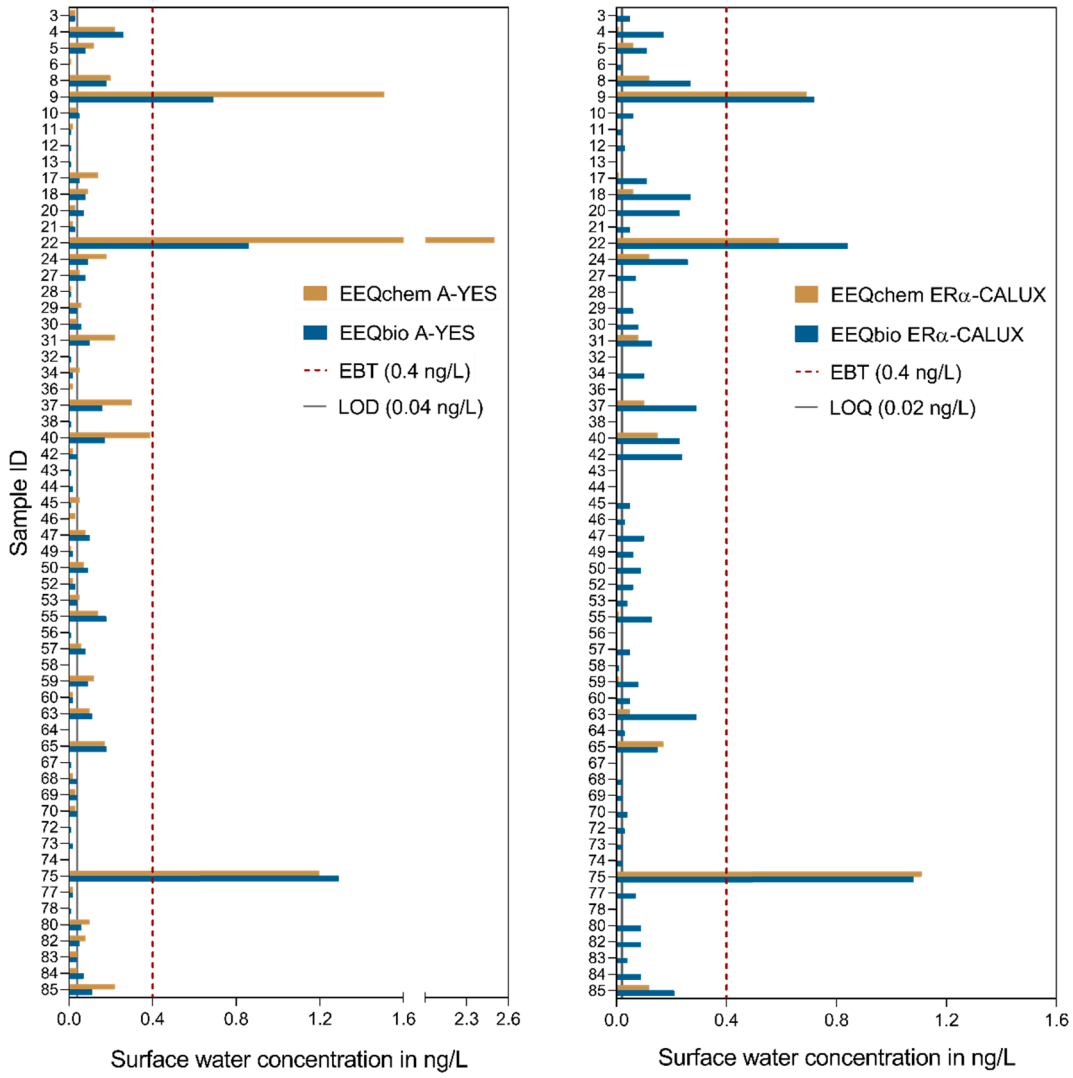


Fig. 1. E2-equivalent concentrations (EEQ, ng/L) of surface water samples with anonymous sample ID measured by ERα-CALUX and A-YES (EEQ_{bio}; blue bars) and calculated based on LC-MS/MS, (EEQ_{chem}; orange bars). Red dashed lines indicate the suggested effect-based trigger value (EBT) of 0.4 ng/L for estrogenicity (Loos, 2012; Kunz et al., 2015; Kase et al., 2018). Samples that were not detected by either method are not shown and the average assay-specific detection (LOD for A-YES) and quantification limits (LOQ for ERα-CALUX) are indicated by grey lines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

over the freeze–thaw process for estrogens and storage at 4 °C for NSAIDs was established in the previous studies (Könemann et al., 2018; Rechsteiner et al., 2020; Frey and Scheurer, 2020).

The estrogenic positive control samples were correctly quantified (i.e., within an acceptable 20% margin of error) by both EBMs and chemical analysis, with one exception. In one of the positive control samples, a slightly higher effect concentration was measured: 0.73 ng EEQ/L instead of 0.6 ng EEQ/L, resulting in a measurement accuracy of 122% (Table S3 in Supplementary data).

The recovery of spiked diclofenac activity for three concentrations varied between 96 and 106%, with an average of 102% in the NSAID assay. This accuracy was proven to be sufficient with a relative standard deviation of 2 – 8% (Table S4, Supplementary data).

3. Results and discussion

3.1. Steroidal estrogen monitoring

3.1.1. Overall low estrogenicity - representing effective estrogenicity reduction efforts in Europe?

In 71 water samples screened, estrogenic activity, expressed as E2-equivalent concentration (EEQ_{bio}), ranged from non-detected (<LOD) to 1.3 ng EEQ/L in A-YES, and ≤ 1.1 ng EEQ/L in ERα-CALUX. LC-MS/MS produced more non-quantifiable results and generally predicted lower estrogenic activity (EEQ_{chem}) than was detected by the bioassays (Fig. 1, Table S5, Supplementary data). Only three samples exceeded the EBT of 0.4 ng EEQ/L, according to both ERα-CALUX and A-YES, as well as based on LC-MS/MS results. Bioassays measured similar estrogenic activities in the samples that exceeded the EBT.

In two samples, higher EEQs were observed by chemical analysis compared to A-YES EEQs. EEQ_{chem} A-YES in sample #9 and #22 were 1.5 and 2.5 ng EEQ/L (Fig. 1) and exceeded the EEQ_{bio} A-YES of 0.7 and 0.9 ng EEQ/L (Figs. 1 and 3). These samples contained elevated concentrations of E1 and E2, above their respective PNEC (Fig. 3). A-YES has an affinity to E1, which is about 20 times higher (relative potency of 0.22) than ERα-CALUX (relative potency of 0.01) (Table S2, Supplementary data). An assay-specific EBT that is optimized for the specific assay sensitivity and specificity – as shown in (Brion et al., 2019) – could help overcome such differences. Based on the individual estrogen concentrations, higher EEQ_{chem} A-YES were expected for samples #9 and #22, but lower activities (EEQ_{bio} A-YES) were actually measured (0.7 and 0.9 ng EEQ/L). This could also be explained by the presence of

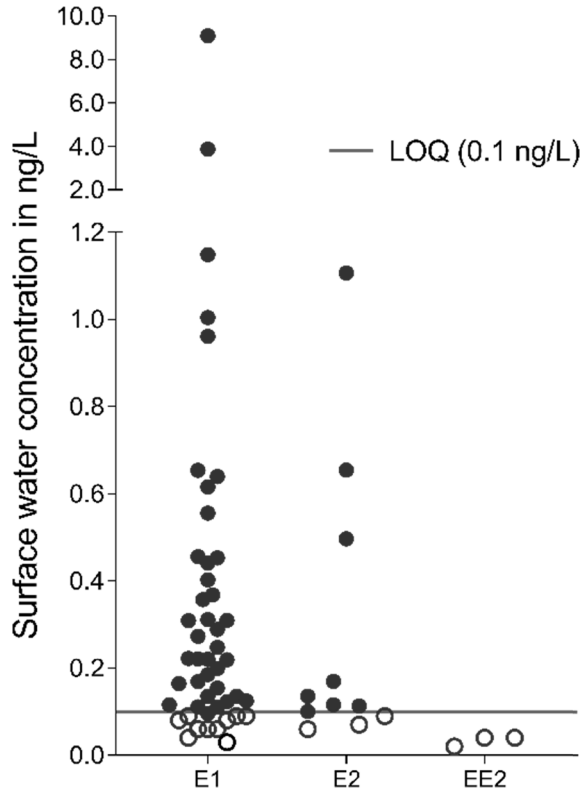


Fig. 2. Concentration distribution of steroidal estrogens (E1, E2 and EE2) in 71 analyzed surface water samples (ng/L). EE2 was detected in three, E2 in 11 and E1 in 47 samples, but many of them were not robustly quantifiable (between LOD and LOQ) presented as open circles. LOQ is depicted by the grey line.

antagonistic substances in the samples, which mask the effect of the agonistic substances, a known phenomenon from other studies (Gehrmann et al., 2018; Itzel et al., 2018, 2020b). However, not observing this in ERα-CALUX suggests the validity of the first explanation (i.e., assay-specific relative potencies of E1 towards E2 causing this difference between EEQ_{chem} and EEQ_{bio} in A-YES).

Overall, estrogenicity detected in this study was relatively low (<2 ng EEQ/L) compared to earlier monitoring studies. In the past, elevated concentrations of natural and exogenous estrogens were detected (e.g., up to 17.9 ng EEQ/L in (Jarošová et al., 2014) and released into effluent receiving surface waters (e.g., up to 4.4 ng EEQ/L in (Thomas et al., 2001). In the last two decades, various efforts have been made to control and eliminate the release of estrogens into aquatic ecosystems. For example, upgrading and optimizing of wastewater treatment plants may have contributed to reducing steroidal estrogen concentrations. Concentrations as low as 6 ng/L were lately found in treated effluents (e.g., Williams et al., 2012) and < 1.5 ng /L in European running waters (e.g., Murk et al., 2002; Vermeirssen et al., 2008; De Baat et al., 2019). In Switzerland, an increasing number of WWTPs are being equipped with ozonation and/or active carbon filtration, which further reduces release of estrogens to surface waters. In treated effluents 0.1–2.6 ng EEQ/L (Schönborn et al., 2017) and in small to medium-sized streams 0.04–0.85 ng EEQ/L, and no exceedances of the EBT set (Kienle et al., 2019) were lately reported. Long-term monitoring and a thorough review of steroidal estrogens in effluents and/or effluent receiving surface waters would, however, be crucial to provide 1) data on trends in estrogenicity and 2) indicate possible causes for such trends.

In about 90% of the investigated samples (65 of 71) individual estrogen concentrations were below 0.8 ng/L (Fig. 2, Table S5, Supplementary data). Only in one sample, a higher E2 concentration was detected. Five samples had an elevated E1 concentration (0.9–9.1 ng/L), but this steroidal estrogen has the lowest estrogenic potential among the three targeted estrogens (Table S2, Supplementary data): five to ten times lower than E2.

The estrogen concentrations found in this study (Fig. 2) fell just within the same low range, although the sampling stations were affected by agricultural runoff and/or discharge from wastewater treatment plants (WWTP) to different degrees ranging from low to high (Supplementary data, Table S1). However, it is important to note that the sites investigated were chosen by the participating countries taking the known or expected presence of measured compounds/effects into account. They were sampled at different times and the monitoring campaign was not designed to interpret spatial differences in pollution levels. For those reasons, no conclusion could be drawn on the impacts

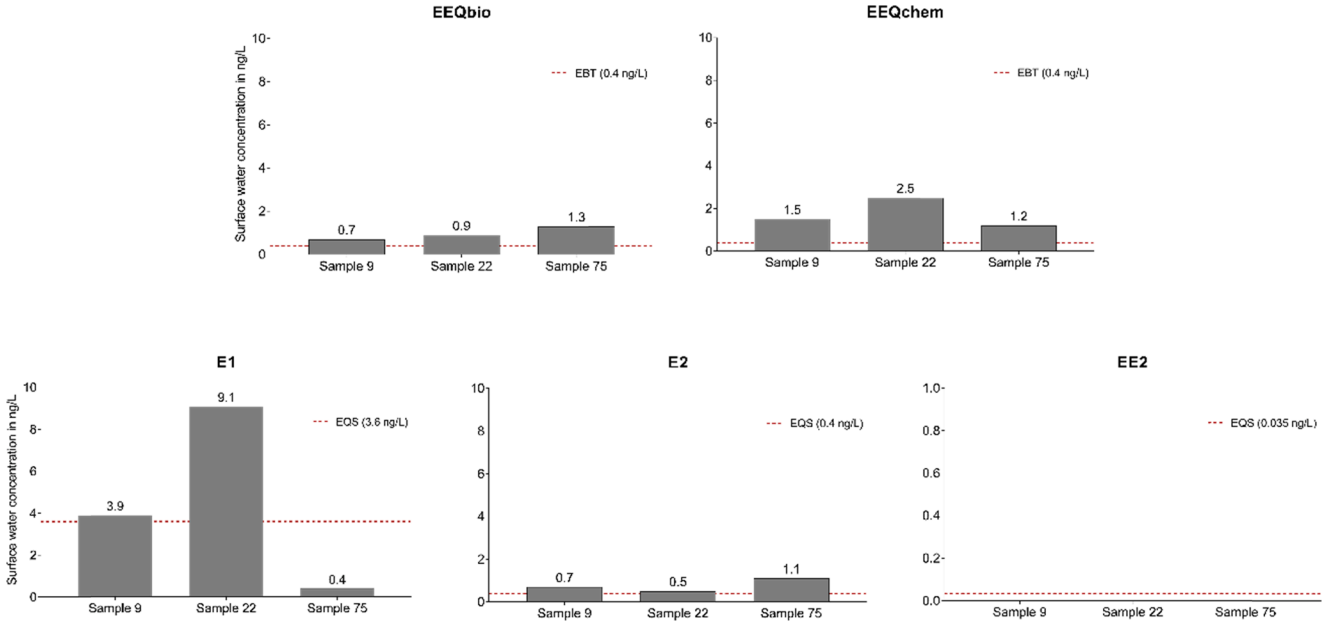


Fig. 3. Comparing measured (EEQ_{bio}) and calculated (EEQ_{chem}) estrogenic activities in A-YES of selected samples (with sample IDs #9, #22, #75) with EBT exceedance (greater than0.4 EEQ ng/L) in the upper part of the figure. The steroidal estrogen (E1, E2 and EE2) composition in these samples showed in the lower part of the figure. EEQ_{chem} are calculated based on the measured individual steroidal estrogens and their relative potencies (Table S2, Supplementary data).

of effluent discharges, agriculture, or other anthropogenic activities.

It is positive, that estrogen pressures are largely below current EBT values, still exceedances did occur, and many samples are within a factor of 10 of current EBT values. However, EBT values may change due to the integration of future effect data or modifications to methods used to derive EBT values. Thus, should new research lead to significant reductions in EBT values (e.g., a reduction by a factor of 10), there would be risk of frequent and widespread EBT value exceedances.

In vitro methods are effective tools to not only identify contaminated sites, but serve as indicators of endocrine disruption in aquatic organisms (Brion, 2019; Sonneveld, 2006). To explore whether the measured *in vitro* estrogenic effects can trigger a response in the zebrafish embryo, a subset of sample extracts (n = 20) was tested using the *in vivo* EASZY assay. No estrogenicity was detected in the field blank. For most of the tested samples, no induction of the GFP driven by the ER-regulated cyp19a1b promoter was measured in the developing brain of zebrafish embryos (Table S6, Supplementary data) and only a few samples were active *in vivo* leading to weak, but significant estrogenic responses observed at the highest sample enrichment tested (Table S6, Supplementary data). The lack of estrogenic responses in EASZY was expected, as *in vitro* estrogenic activity was low in most samples, and below the assay-specific EBT value established for ERα-CALUX, i.e., 0.28 ng/L (Brion et al., 2019) or below the EBT established for estrogenicity in general (0.4 ng/L). However, for sample #65, a mismatch was noticed. This sample induced a weak, but significant GFP induction *in vivo,* while low estrogenic activity was detected in both ERα-CALUX (0.15 ng EEQ/L) and A-YES (0.18 ng EEQ / L) (Table S6 and Figure S2, Supplementary data). Similarly, no *in vivo* response was expected based on chemical analysis of the three steroidal estrogens. The presence of unidentified substances which are active *in vivo* but not *in vitro* could explain this discrepancy. It has been shown, for instance, that some chemicals required metabolic activation before eliciting an estrogenic response leading to ER-dependent induction of brain aromatase in EASZY (Brion et al., 2012; Cano-Nicolau et al., 2016). Such differences between *in vitro* assays and *in vivo* responses may reflect the pharmacodynamic interactions of the endocrine active substances in embryos.

In vivo estrogenic activities were indeed measured at most sites for which *in vitro* estrogenic activity were above the EBT (Brion et al., 2019) confirming the risk posed by environmental estrogens at these locations. Intriguingly, no *in vivo* estrogenic effect was measured for site #75 while this sample has a similar E1 and E2 composition to sample #9 and #22, two sites showing significant *in vivo* GFP induction (Table S6 and Figure S2, Supplementary data). The concentration–response curve of sample #75 was atypical as a difference was noticed between the highest and the lowest REFs suggesting that compounds present in the organic extract could have negatively interfered with the steroidal estrogens (Cheshenko et al., 2007; Serra et al., 2018).

Altogether, the comparison of the *in vivo* and *in vitro* responses confirmed the relevance of an EBT value to identify polluted sites based on *in vitro* measurement of estrogenic activity. Similarly, high sensitivity (true positive rate) and specificity (true negative rate) were shown for both *in vitro* assays (100% and 86% respectively, Table S6, Supplementary data) meaning that there is a high predictive power of using *in vitro* estrogenic activities to assess the quality of water bodies with no or low probability of false positive and negative assessments.

3.1.2. Do bioassays outperform target chemical analysis?

Samples with higher estrogenic (effect) concentrations (i.e., above the threshold) were correctly identified with both chemical analytical and effect-based methods. Discrepancies were primarily seen for samples with low or near-threshold levels. The bioassays were able to quantify estrogenicity in the majority of samples (51 samples in ERα-CALUX and 55 in A-YES, Table 1, Table S5 in Supplementary data), while LC-MS/MS showed more non-quantifiable results and therefore made an estrogenic effect estimation below EBT difficult (Table 1, Table S5 Supplementary data). This was most likely due to the fact that

Table 1
Number of samples with limit value exceedances, non-quantifiable and quantifiable results of the total number of 71 surface water samples. Quantifiable results (>LOQ) are shown with and without standard addition. Standard addition was performed by adding the respective LOQ concentration to the surface water samples to detect estrogen concentrations below or near the method LOQ (0.1 ng/L for E1, E2, EE2). In the last column, samples with concentrations < 0.1 ng/L were also categorized as quantified samples, as after standard addition those low concentrations could be detected. Standard addition was only performed for chemical analysis and not for the bioassays. Therefore, no EEQ_{bio} ERα-CALUX was determined, as indicated by "-".

Nr. of samples	Exceedances (≥PNEC or EBT)	Non-detects (<LOQ)	Quantified results (>LOQ)	Quantified results after standard addition
EEQbio ERα-CALUX	3	20	**51**	–
EEQchem ERα-CALUX	3	54	17	**47**
E1	3	33	38	67
E2	3	63	8	46
EE2	(2)	71	0	57

LOQs of LC-MS/MS are about ten times higher (~0.1 ng/L for each steroidal estrogen targeted) than those of the EBMs (~0.01 ng EEQ/L). Other studies have shown that bioassays are sensitive enough to detect estrogenicity even in drinking waters containing only trace levels of steroidal hormones and phenolic compounds (e.g., Leusch et al., 2017).

With optimal extraction of the water sample, organic micropollutants can be separated from other matrix components and higher sensitivity of both effect-based and chemical analytical methods (i.e., lower method detection and quantification limits) can be achieved (Brack et al., 2019). However, excessive enrichment can interfere with the subsequent analyses (Simon et al., 2015). Co-extracted sample matrix can cause cell death or inhibit cell growth (cytotoxicity) in a bioassay or induce co-eluting peaks or ion suppression in chemical analysis. The latter (ion suppression up to 70% with average of 40%) was an issue in the majority of samples, while no matrix related interferences were observed in the effect-based analyses. The difference between the relative enrichment factors, at which the samples were tested in the analytical and effect-based methods, has to be noted, when discussing the phenomenon of matrix effect and considered as the reason for the observed differences. For LC-MS/MS, the 1000-fold enriched sample extracts were further concentrated by a factor of four (i.e., $REF_{LC\text{-}MS/MS} = 4000$). For the bioassays the extracts were diluted prior the testing: $REF_{CALUX} = 20$ and $REF_{A\text{-}YES} = 500$.

When making further comparisons, a crucial difference between the two screening tools must be discussed. Bioassays measure total biological effect of all substances with similar mode of action simultaneously present, even of those not targeted or not known to appear in the water extract (e.g., emerging contaminants, transformation products), and account for their mixture effects. Chemicals in the sample act in concert and can enhance or mask each other's toxic effect (Brack et al., 2019). Despite the numerous advantages, they cannot differentiate between the chemicals present and reveal their identity, unless they are combined with chromatographic techniques (e.g., thin-layer chromatography, (Buchinger et al., 2013; Bergmann et al., 2020; Moscovici et al., 2020; Baetz et al., 2021). Analytical tools, LC-MS/MS in this case, on the other hand accurately identify and quantify the targeted analytes, which is of great importance in terms of monitoring prioritized substances like those on the WFD Watch List. All things considered, determination of the total *in vitro* effect and hence a link to possible *in vivo* adverse outcomes is of higher importance for the sake of biological relevance and risk assessment of hazards, than the targeted identification of elected chemicals. The analysis of other targeted phenolic compounds

(bisphenol-A, BPA and octylphenol, OP) is an example thereof. Elevated BPA and OP concentrations (up to 25 ng/L) were found in the majority of our samples without any analytical detection challenges, yet their contribution to the total biological effect was negligible (<5%; data not shown) due to their low estrogenic potency (e.g., BPA is ~ 50.000 times less bioactive than E2, (Murk et al., 2002; Simon et al., 2019).

Along this line of reasoning, "effect-based methods are key", but their combined use with chemical analysis allows the most comprehensive water quality assessment strategy (Brack et al., 2019). The current study also proved the good screening potential of EBMs and thereby their applicability as a diagnostic and early-warning tool. Chemical analysis can then identify spills, potent risk drivers, and emerging contaminants and reveal contamination trends at priority sites (e.g., Wernersson et al., 2015; van der Oost et al., 2017b; Brack et al., 2019; De Baat et al., 2019; Simon et al., 2019).

3.1.3. Options to tackle analytical detection challenges

LC-MS/MS showed in various studies to have difficulty detecting low (pg/L) estrogenic concentrations (this study and e.g., Loos, 2012; Kunz et al., 2015; Kase et al., 2018; Könemann et al., 2018). Studies reported various options to overcome this challenge.

(1) Use sensitive *in vitro* estrogenic screening assays, as early-warning tools to identify samples with estrogenic potency. Their applicability has been shown in this and other studies (Leusch et al., 2010; Könemann et al., 2018; Hettwer et al., 2018).
(2) Assess exposures and possible river concentrations through modeling based on (primarily human) consumption data (Johnson et al., 2013).
(3) Improve sample preparation by extensive clean-up (e.g., online SPE) to eliminate co-extracted interfering matrix components (Goeury et al., 2019; Itzel et al., 2020a).
(4) Choose another adjusted analytical method, such as GC–MS/MS after derivatization of the extracts to reach low detection limits for E2 and EE2 (10 pg/L = 0.01 ng/L) (Zacs et al., 2016; Itzel et al., 2020a).
(5) Replace non-detects with the target analyte specific LOD or LOD/2 concentration instead of zeroing them. This constitutes a conservative approach to avoid underestimating risk (Kase et al., 2018). A similar concept is applied in the EU food regulation of dioxins and dioxin-like compounds (EU, Regulation (EU) No 589/, 2014).
(6) Apply the quantitative analysis approach of "standard addition" (represented in Figure S1). Spiking the aliquot of the sample with a known amount of target analyte concentration to improve quantification originally masked by the matrix (Tavazzi et al., 2016) or reach trace levels of steroidal estrogens (Cimetiere et al., 2013).

Although bioassays (1) hold considerable promise to enhance exposure and risk assessments, their regulatory acceptance is not yet secured. Modeling approaches (2) have certain limitations if data are not fully available. They are not precise enough as removal rates may have high uncertainty. Water laboratories often choose option (3) and (4) to improve their current analytical techniques over implementing or expanding their facilities with alternative effect-based technologies (1). As we had no possibility to explore options (2) to (4) and option (5) was extensively investigated in our previous study (Kase et al., 2018), we decided to additionally look into the standard addition approach (option 6) to explore detections in the range of the LOQ of our LC-MS/MS method.

Table 1 shows that standard addition allowed for accurate quantification of E2 and EE2 in the surface water samples and compensated for matrix effects. After standard addition, EE2 could be detected in 57 samples, E2 in 46 and E1 in 67 (Table 1, Figure S4 in Supplementary data). The standard addition exercise did not only lead to accurately quantifiable low bioactive steroidal estrogen concentrations (Figure S3 and S4, Supplementary data), but also to an excellent agreement between ERα-CALUX and LC-MS/MS (Table 1, detailed data are not shown), with 51 and 47 quantified samples. This confirms the sensitivity, reliability, and good screening potential of the EBMs at critical, biologically relevant concentrations and points out no need for such additional experiment in routine LC-MS/MS analysis to tackle the detection challenges for E2 and EE2.

3.1.4. Practical considerations: Sample filtration, enrichment, in vitro assay selection and data interpretation

WFD requires analyzing the whole water phase including both dissolved and suspended material-bound contaminants (Richter et al., 2016; Yarahmadi et al., 2018; EU GD Nr. 27, 2011; Loos et al., 2018a). Filtration may only lead to loss of undissolved, very hydrophobic contaminants. For compounds below a Log K_{ow} of 6 (thus diclofenac, E1, E2 and EE2 in this case) discrepancies between total and dissolved concentrations do not become evident (Richter et al., 2016; Loos et al., 2018a; Yarahmadi et al., 2018; EU GD Nr. 27, 2011). In our previous study, we investigated the impact of filtration (using glass fiber filters with 2.7 μm retention) on the recovery of the targeted steroidal estrogens. No altered estrogenic concentrations were found after filtration in the control or the tested water samples (Könemann et al., 2018). Walker and Watson disclosed no significant sorption losses of E1, E2 or EE2 either, with even smaller glass fiber filters (1.0 and 0.3 μm) (Walker and Watson, 2010). Although they used relatively high estrogen concentrations and warranted further investigation at trace levels.

There is another important consideration for the analysis of water samples. The current effect-based and chemical analyses of steroidal estrogens were conducted with enriched samples. This was necessary to achieve low detection limits and appropriate sensitivity. Any sample preparation step is, however, a manipulation step and predetermines the analytes present in the extract (Simon et al., 2015; Neale et al., 2018; Brack et al., 2019). To capture all contaminants that are present in a water sample, its direct testing (i.e., without filtration, enrichment or clean-up) would be optimal. But this might not be feasible for every *in vitro* assay and analytical method. The validation data of ERα-CALUX and A-YES, however, showed in a previous study (e.g., Gehrmann et al., 2018), that the used enrichment protocol does not affect the estrogenic activity of the whole sample. In case of any concerns about very polar, non-enrichable estrogenic active compounds both ERα-CALUX and A-YES assays can robustly be performed with native water samples (ISO 19040-2:2018, Water quality; ISO 19040-3:2018, Water quality). The detection limits for direct testing will, however, be higher than those for sample extracts testing and increase to ~ 0.5 ng EEQ/L for ERα-CALUX assay and on average 1.8 ng EEQ/L for A-YES (Hettwer et al., 2018; Rechsteiner et al., 2020; ISO 19040-2:2018, Water quality; ISO 19040-3:2018, Water quality).

Before deciding on extract or direct testing, the best suitable *in vitro* method has to be selected for the assessment of estrogenic activities from the numerous alternatives currently available (e.g., GWRC Reports 2006 and 2008; Leusch et al., 2010; Kunz et al., 2017). Besides technical and assay parameters (such as sensitivity, robustness, accuracy), ease of use and interpretation, turnaround time, required equipment and consumable costs, availability of technical support and kit-based systems are of all high relevance to be considered. Those selection criteria are extensively discussed in other projects and studies (Schriks et al., 2015; Leusch et al., 2018). Here, excellent agreement was found in the predictive and quantification capacity of ERα-CALUX and A-YES. Estrogenic equivalences, non-detects and exceedances were correctly identified by both assays, despite the differences in their performance protocol, assay characteristics and relative potency towards estrone (E1) in the bioassays. The A-YES makes use of yeast cells and tested extracts re-dissolved in water at a relative enrichment of 500. ERα-CALUX is a human cell line-based assay and carried out with extracts re-dissolved in DMSO at relative enrichment of 20. A-YES has an about 20 times higher

affinity to E1 (REP = 0.22), than ERα-CALUX (REP = 0.01). Since predominantly E1 was found in the surface water samples, A-YES showed generally higher E2-equivalences, but trigger value exceedances were correctly identified by both methods.

To identify exceedances and compare methods, we used a straightforward limit value approach. Individual steroidal estrogen concentrations (determined by LC-MS/MS) were compared to their respective PNEC values. Effect-based concentrations (both EEQ_{bio} and EEQ_{chem}, i.e., the measured and LC-MS/MS-based calculated effects) were compared to the EBT of 0.4 ng EEQ/L derived based on the PNEC of the bioassays' reference chemical, E2 (0.4 ng/L) (Loos, 2012; Kunz et al., 2015). Thereby, we used limit values that both rely on PNEC levels and enabled a simple method comparison.

It is important to note that various limit value proposals are currently available based on statistical, theoretical and practical derivation approaches (Besselink et al., 2017). EBTs can be assay-specific addressing differences in relative potencies towards the known effect-driving chemicals (Escher et al., 2018; EU Proposal CIS WG, 2020) and *in vitro/in vivo* correlations (Brion et al., 2019). EBTs can also be derived based on biological equivalent concentrations from existing regulated chemical guideline values and published biological effect concentrations for these regulated chemicals (Escher et al., 2018). Besselink *et al.* derived EBT, for instance, based on the estrogenic activities of a large set of water samples (n = 1000), and used a limit value, below which a certain percentage of the analysis results fall (Besselink et al., 2017). The percentage chosen depended on the sample size. Eventually, a scenario must be chosen that is solid and fits the purpose of the study, and do not over- or underestimate the risk by choosing a trigger value that is either too low or too high. We wanted to point out these differences and note that the selection of a limit value (EBT), predetermines the ecological-based risk assessment and the number of exceedances. The currently reported and available EBT's for estrogenic activity, however, fall in a narrow range of 0.1 and 0.8 ng/L. Despite the fairly diverse EBT derivation methods this rather small difference can be considered highly acceptable.

3.2. Diclofenac monitoring

A subset of samples (n = 39) was investigated for COX-inhibitors in a similar approach to the estrogen monitoring. The NSAID *in vitro* assay with a detection limit of 6 ng/L diclofenac fulfilled the required maximum acceptable method detection limit (MDL) for the EU Watch List monitoring (i.e., smaller to the PNEC of 50 ng/L; (EU, Commission Implementing Decision (EU), 2015; Loos et al., 2018b). Diclofenac equivalents (DicEQs) of 12 of 39 samples (31%) exceeded the suggested predicted no-effect concentration (PNEC = 50 ng/L, (Loos et al., 2018b) according to both the EBM and the chemical analysis (Fig. 4, Table S7 in Supplementary data). In the majority of the samples, predominantly diclofenac was detected (Table S7, Supplementary data). The other targeted COX-inhibitors ibuprofen, naproxen and salicylate for example were found in some samples (15 of 39), but only in minor concentrations (up to 96, 100 and 32 ng/L respectively). Therefore, it can be stated that almost exclusively diclofenac is responsible for the measured biological effects, i.e., DicEQs. Ten samples were below the limit of detection of LC-MS/MS and two samples below the limit of detection of the EBM. No matrix effect was noticed during LC-MS/MS quantification or observed *in vitro*. In sample #63, with the highest diclofenac activity and concentration, high organic load was observed indicating higher ratio of treated wastewater in the receiving surface water sampled. Extraction and analysis were repeated for this sample and the result thereof was reported after similar outcomes were found.

Altogether, the *in vitro* biosensor cell line gave highly comparable results to LC-MS/MS and indicated diclofenac concentrations in the same range (12–223 ng/L with the *in vitro* assay and 11–310 ng/L with LC-MS/MS). With one exception (sample #63 with 1.1 µg diclofenac/L), the samples contained diclofenac up to ~ 300 ng/L showing a similar concentration pattern for European surface waters, as recently reviewed by Satishkumar and co-workers (Sathishkumar et al., 2020). According to this review, diclofenac concentrations can vary from country to country and in different surface water types. The highest concentrations were found in rivers at densely populated urbanized areas and in streams, because of intensive farming. The current diclofenac data set

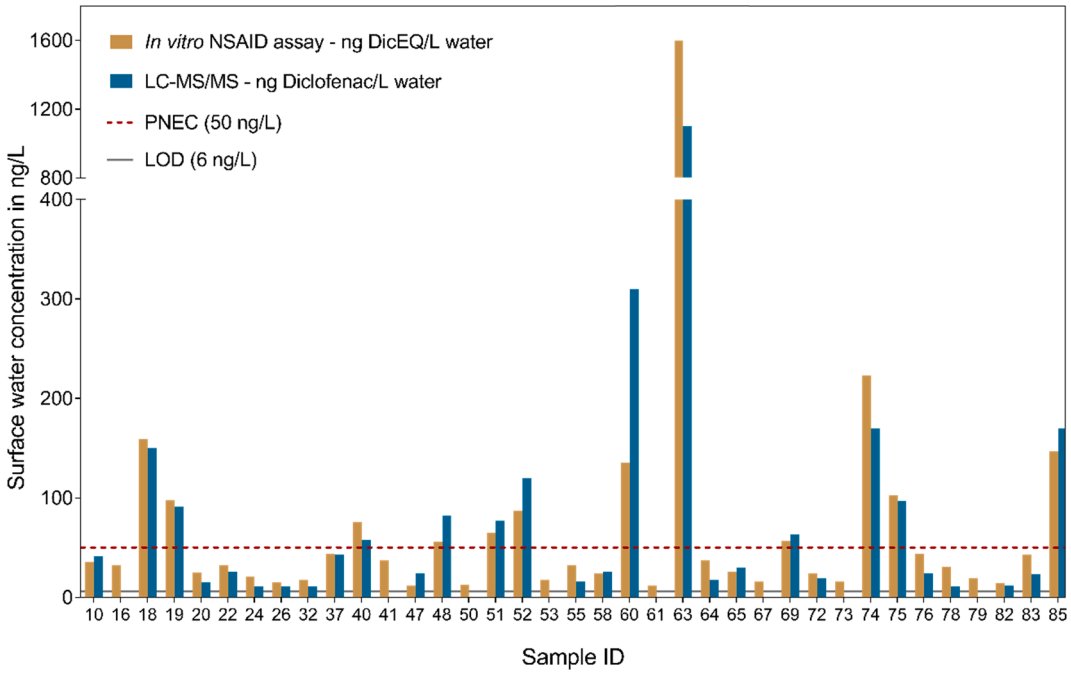


Fig. 4. Surface water samples analyzed by chemical analysis (LC-MS/MS, blue bars) and the *in vitro* NSAID assay (orange bars). Chemical and biological effect concentrations were compared to the predicted no-effect concentration (PNEC) of 50 ng/L for diclofenac represented by the red dashed line to identify exceedances. Grey line represents the LOD of the *in vitro* NSAID assay. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

point out that many sampled sites had WWTP effluent or agricultural input.

In vitro measured effect concentrations were in most of the samples slightly higher than the diclofenac concentration, except in sample #60. Here, chemical analysis was significantly higher (310 ng Diclofenac/L) than the bioassay (135 ng DicEQ/L). One explanation may be the presence of antagonists, compounds that can block or reduce the effect of the agonistic counterparts. This emphasizes the predictive power of EBMs for biological relevance by accounting for mixture effects. Another explanation could be a reduced recovery of diclofenac on the SPE column. Reduced recovery would produce a lower DicEQ in the *in vitro* assay, whereas for chemical analysis the internal standard corrects for recovery losses. The elevated standard deviation (65%) in the *in vitro* assay for this sample may support the co-extracted "impurity" causing difficulties in the cellular test system. To avoid a loss of NSAIDs due to reduced extraction efficiency, and to ensure the measurement of the complete complex mixture, an alternative vacuum concentration was developed and validated (data not shown, but reported in another study (Frey and Scheurer, 2020). For future COX-inhibition analyses this method is recommended over SPE.

Although LC-MS/MS and the *in vitro* NSAID bioassay produced comparable data in the set of surface waters analyzed, the application of the *in vitro* technique as a prescreening tool is still of high relevance. We showed that diclofenac primarily accounted for the measured effects, but the composition of other surface waters (collected at different locations at different time points) may vary.

Important to note that neither the chemical nor the effect-based analysis could be performed in the same extracts prepared for the estrogen monitoring. A suitable sample preparation had to be developed and applied to the COX-inhibitors. The lack of a generic sample preparation method suitable for the investigation of a broad range of toxic endpoints and target analytes is a downside in environmental water quality assessment, regardless of the monitoring method used (i.e., effect-based or chemical).

4. Conclusions and outlook

High-quality monitoring data were produced and the applicability of effect-based methods (EBMs) was shown for the monitoring of specific groups of substances in the aquatic environment under the EU WFD.

The applied EBMs of steroidal estrogens were more sensitive having fewer non-quantified samples (28%) than chemical analysis (76%), pointing out their excellent potential as screening tools. At the same time, it also implies that analytical methods are faced with detection challenges of the target analytes at low concentrations (quantification of EE2 at its PNEC level remained a hurdle for conventional LC-MS/MS). According to current knowledge and current limit value (EBT) approaches, the estrogenic concentrations (predominantly < 1 ng/L) found in the sampled surface waters are low and pose low risk to aquatic organisms.

Analytical and effect-based methods provided concordant results for the COX-inhibitors, as no detection difficulties nor other COX-inhibitors than diclofenac were identified at elevated concentrations in the samples. The presence of diclofenac at the majority of the analyzed sampling sites indicates WWTP effluent discharges in those surface waters.

In the view of the results, a combined use of analytical and effect-based methods is recommended to get the benefits of both and conduct a comprehensive water quality assessment to:

- monitor specific groups of substances in the aquatic environment,
- receive indication on substances and their contribution to the measured effects and prioritise risk drivers,
- distinguish between contaminated and clean sites, and conduct further in-depth analysis at hotspots, while avoiding costly analysis of sites with no risks,
- EBMs appeared to quantify effects in more samples than LC-MS/MS, but they cannot distuinguish the effects of chemicals causing the effects,
- exclude false negative measurements.

Future studies should define necessary framework conditions for the implementation of EBMs into regulatory monitoring programs to enable a paradigm shift in current monitoring practices towards effect-based screening. Updating the list of routinely monitored chemicals under the WFD should give the possibility to not only include emerging contaminants, but also embed novel approaches, such as EBMs, into monitoring these contaminants.

CRediT authorship contribution statement

Eszter Simon: Supervision, Project administration, Visualization, Formal analysis, Writing – original draft. **Anja Duffek:** Conceptualization, Methodology, Project administration. **Cordula Stahl:** Methodology, Investigation. **Manfred Frey:** Methodology, Investigation. **Marco Scheurer:** Investigation. **Jochen Tuerk:** Investigation. **Linda Gehrmann:** Investigation. **Sarah Könemann:** Visualization. **Kees Swart:** Investigation. **Peter Behnisch:** Investigation. **Daniel Olbrich:** Investigation. **François Brion:** Investigation. **Selim Aït-Aïssa:** Investigation. **Robert Pasanen-Kase:** Conceptualization, Methodology, Funding acquisition, Project administration. **Inge Werner:** Resources. **Etienne Vermeirssen:** Validation, Resources. **All authors:** Reviewing and Editing.

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.

Acknowledgement

The authors thank all partners for their in-kind collaboration, interest and providing the surface water samples for analysis and appreciating the relevance and importance of this project. Special thank goes to the participating countries and their partner in water sampling in alphabetic order: Synlab Umweltinstitut GmbH (AT), Cellule écotoxicolgie (BE), Amt für Wasser und Abfall Bern (CH), Amt für Wasser und Energie St. Gallen (CH), AWEL Zürich (CH), Amt für Umwelt, Thurgau (CH), The Danish Environmental Protection Agency (DK), Niedersächsischer Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz (DE), Bayerisches Landesamt für Umwelt (DE), LLBB Landeslabor Berlin Brandenburg (DE), Landesamt für Umwelt Rheinland-Pfalz (DE), Finnish Environment Institute (FI), AQUAREF (FR), Latvian Environment Geology and Meteorology Centre (LV), Rijkswaterstaat Laboratorium (NL), Scottish Environment Protection Agency (SCO), Umeå University (SE), Water Research Institute VUVH Bratislava (SK), Slovenian Environment Agency (SL).

The authors are grateful to Andrea Schifferli for her technical assistance at the Swiss Centre for Applied Ecotoxicology (Switzerland). We are also grateful to Mario Carere (National Institute of Health, Rome, Italy) and Cornelia Kienle (Swiss Centre for Applied Ecotoxicology, Switzerland), who supported the concept and visualization of the study. We thank Marina Ricci (EU JRC-Joint Research Centre) for the fruitful discussions on sample preparation and WFD requirements.

Funding

This research did not receive any specific grant from agencies in the public, commercial, or not-for-profit sectors. The project was initiated and funded to a great extent by the Swiss Centre for Applied Ecotoxicology. This included the costs of the planning and organization of the

project, sample transports, sample workup for steroidal estrogen analyses, LC-MS/MS and ERα-CALUX analyses of the extracts. Partners collected the water samples at their own costs. The diclofenac analyses, as part of the project "EffPharm - Validation of the NSAID in vitro assay for biomonitoring of NSAID activities in surface waters", were funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (project no. 3717634400). IUTA (DE) funded the A-YES estrogenicity analysis of the samples and new diagnostic GmbH, Berlin (DE) supported the study by providing yeast cells. INERIS (FR) received a funding from AQUAREF as part of the scientific and technical program for the year 2019, theme "New tools and knowledge to optimize monitoring strategies" and the DRC60 from the French Ministry of Environment to conduct the *in vivo* EASZY assay.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2021.107033.

References

Adeel, M., et al., 2017. Environmental impact of estrogens on human, animal and plant life: A critical review. Environ. Int. 99, 107–119.
Altenburger, R., et al., 2015. Future water quality monitoring — Adapting tools to deal with mixtures of pollutants in water resource management. Sci. Total Environ. 512–513, 540–551.
Baetz, N., et al., 2021. High-performance thin-layer chromatography in combination with a yeast-based multi-effect bioassay to determine endocrine effects in environmental samples. Anal. Bioanal. Chem. 413 (5), 1321–1335.
Bergmann, A.J., et al., 2020. Estrogenic activity of food contact materials—evaluation of 20 chemicals using a yeast estrogen screen on HPTLC or 96-well plates. Analyt. Bioanalyt. Chem. 412 (19), 4527–4536.
Bernhard, K., et al., 2017. Two novel real time cell-based assays quantify beta-blocker and NSAID specific effects in effluents of municipal wastewater treatment plants. Water Res. 115, 74–83.
Besselink, H., et al., 2017. Validation and regulatory acceptance of bio-based approaches to assure feedstock, water & product quality in a bio-based economy. Indus. Crops Prod. 106, 138–145.
Brack, W., 2015. The Challenge: Prioritization of emerging pollutants. Environ. Toxicol. Chem., 34, 10, 2181–2181.
Brack, W., et al., 2016. Effect-directed analysis supporting monitoring of aquatic environments — An in-depth overview. Sci. Total Environ. 544, 1073–1118.
Brack, W., et al., 2017. Towards the review of the European Union Water Framework Directive: Recommendations for more efficient assessment and management of chemical contamination in European surface water resources. Sci. Total Environ. 576, 720–737.
Brack, W., et al., 2019. Effect-based methods are key. The European Collaborative Project SOLUTIONS recommends integrating effect-based methods for diagnosis and monitoring of water quality. Environ. Sci. Eur. 31 (1), 10.
Brion, F., et al., 2012. Screening estrogenic activities of chemicals or mixtures in vivo using transgenic (cyp19a1b-GFP) zebrafish embryos. PloS one, 7, 5, e36069–e36069.
Brion, F., 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.
Buchinger, S., et al., 2013. Direct Coupling of Thin-Layer Chromatography with a Bioassay for the Detection of Estrogenic Compounds: Applications for Effect-Directed Analysis. Anal. Chem. 85 (15), 7248–7256.
Busch, W., et al., 2016. Micropollutants in European rivers: A mode of action survey to support the development of effect-based tools for water monitoring. Environ. Toxicol. Chem. 35 (8), 1887–1899.
Cano-Nicolau, J., et al., 2016. Several synthetic progestins disrupt the glial cell specific-brain aromatase expression in developing zebra fish. Toxicol. Appl. Pharmacol. 305, 12–21.
Carere, M., P.S., Kase, R., Gawlik, B.M., 2015. The Emerging Contaminants in the Context of the EU Water Framework Directive. In: Petrovic, M., Sabater, S., Elosegi, A., Barceló, D. (Eds.). Emerging Contaminants in River Ecosystems. The Handbook of Environmental Chemistry, vol. 46. Springer, Cham.
Cheshenko, K., et al., 2007. Expression of Zebra Fish Aromatase cyp19a and cyp19b Genes in Response to the Ligands of Estrogen Receptor and Aryl Hydrocarbon Receptor. Toxicol. Sci. 96 (2), 255–267.
Cimetiere, N., et al., 2013. Standard addition method for the determination of pharmaceutical residues in drinking water by SPE–LC–MS/MS. Environ. Technol. 34 (22), 3031–3041.
Combes, R.D., Gaunt, I., Balls, M., 2004. A scientific and animal welfare assessment of the OECD Health Effects Test Guidelines for the safety testing of chemicals under the European Union REACH system. Altern. Lab. Anim. 32 (3), 163–208.
Connon, R.E., Geist, J., Werner, I., 2012. Effect-based tools for monitoring and predicting the ecotoxicological effects of chemicals in the aquatic environment. Sensors (Basel, Switzerland) 12 (9), 12741–12771.
De Baat, M.L., et al., 2019. Effect-based nationwide surface water quality assessment to identify ecotoxicological risks. Water Res. 159, 434–443.
Escher, B.I., 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. 628–629, 748–765.
EU, 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 pursuant to Directive 2008/105/EC of the European Parliament and of the Council (notified under document C(2015) 1756) E. Commission, Editor. 2015.
EU, 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 pursuant to Directive 2008/105/EC of the European Parliament and of the Council and repealing Commission Implementing Decision (EU) 2015/495. 2018.
EU, 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. 2000.
EU, Directive 2013/39/EU of the European Parliament and of the Council of 12 August 2013 amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. 2013.
EU, Guidance Document No. 27; Technical Guidance For Deriving Environmental Quality Standards. 2011.
EU, Proposal for Effect-Based Monitoring and Assessment in the Water Framework Directive. Report to the CIS WG Chemicals on the outcome of the work performed in the subgroup on effect-based methods (EBMs). 2020.
EU, Regulation (EU) No 589/2014 of 2 June 2014 laying down methods of sampling and analysis for the control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation (EU) No 252/2012. 2014.
Frey, M., Scheurer, M., 2020. Validation of the NSAID in vitro assay for biomonitoring of NSAID activities in surface waters (Project No. 3717 63 440 0). German Environment Agency.
Gan, T.J., 2010. Diclofenac: an update on its mechanism of action and safety profile. Curr. Med. Res. Opin. 26 (7), 1715–1731.
Gaulke, L.S., et al., 2009. Estrogen Biodegradation Kinetics and Estrogenic Activity Reduction for Two Biological Wastewater Treatment Methods. Environ. Sci. Technol. 43 (18), 7111–7116.
Gehrmann, L., et al., 2018. (Anti-)estrogenic and (anti-)androgenic effects in wastewater during advanced treatment: comparison of three in vitro bioassays. Environ. Sci. Pollut. Res. 25 (5), 4094–4104.
Goeury, K., et al., 2019. Analysis of Environmental Protection Agency priority endocrine disruptor hormones and bisphenol A in tap, surface and wastewater by online concentration liquid chromatography tandem mass spectrometry. Chromatogr A 1591, 87–98.
GWRC, 2008. Tools to detect estrogenicity in environmental waters – Final report. Global Water Research Coalition (GWRC)/Water Environment Research Foundation (WERF) report. 74pp.
GWRC, 2006. In vitro bioassays to detect estrogenic activity in environmental waters – A literature review. Global Water Research Coalition (GWRC)/Water Environment Research Foundation (WERF) report. 76pp.
Hamers, T., et al., 2013. Expert opinion on toxicity profiling–report from a NORMAN expert group meeting. Integr. Environ. Assess. Manage. 9 (2), 185–191.
Hettwer, K., et al., 2018. Validation of Arxula Yeast Estrogen Screen assay for detection of estrogenic activity in water samples: Results of an international interlaboratory study. Sci. Total Environ. 621, 612–625.
ISO 19040-1:2018, Water quality — Determination of the estrogenic potential of water and waste water — Part 1: Yeast estrogen screen (Saccharomyces cerevisiae). 2018.
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). 2018.
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. 2018.
Itzel, F., et al., 2017. Investigation of full-scale ozonation at a municipal wastewater treatment plant using a toxicity-based evaluation concept. J. Toxicol. Environ. Health, Part A 80 (23–24), 1242–1258.
Itzel, F., et al., 2018. Comprehensive analysis of antagonistic endocrine activity during ozone treatment of hospital wastewater. Sci. Total Environ. 624, 1443–1454.
Itzel, F., et al., 2020a. Ultrasensitive Bestimmung von Hormonen. Chrom+Food FORUM 10, 48–49.
Itzel, F., et al., 2020b. Evaluation of a biological post-treatment after full-scale ozonation at a municipal wastewater treatment plant. Water Res. 170, 115316.
Jarošová, B., et al., 2014. Europe-wide survey of estrogenicity in wastewater treatment plant effluents: the need for the effect-based monitoring. Environ. Sci. Pollut. Res. 21 (18), 10970–10982.
Jobling, S., et al., 2006. Predicted exposures to steroid estrogens in U.K. rivers correlate with widespread sexual disruption in wild fish populations. Environ. Health Perspect. 114 (Suppl 1), 32–39.
Johnson, A.C., et al., 2013. Do concentrations of ethinylestradiol, estradiol, and diclofenac in European rivers exceed proposed EU environmental quality standards? Environ. Sci. Technol. 47 (21), 12297–12304.
Kase, R., et al., 2018. Screening and risk management solutions for steroidal estrogens in surface and wastewater. TrAC Trends Analyt. Chem. 102, 343–358.
Kidd, K.A., et al., 2007. Collapse of a fish population after exposure to a synthetic estrogen. Proc. Natl. Acad. Sci. 104 (21), 8897–8901.
Kienle, C., et al., 2019. Effects of treated wastewater on the ecotoxicity of small streams - Unravelling the contribution of chemicals causing effects. PloS one 14 (12), e0226278–e0226278.

Könemann, S., et al., 2018. Effect-based and chemical analytical methods to monitor estrogens under the European Water Framework Directive. TrAC Trends Analyt. Chem. 102, 225–235.
Kunz, P.Y., et al., 2015. In vitro bioassays to screen for endocrine active pharmaceuticals in surface and waste waters. J. Pharmaceut. Biomed. Anal. 106, 107–115.
Kunz, P.Y., et al., 2017. Effect-based tools for monitoring estrogenic mixtures: Evaluation of five in vitro bioassays. Water Res. 110, 378–388.
Leusch, F.D.L., et al., 2010. Comparison of five in vitro bioassays to measure estrogenic activity in environmental waters. Environ. Sci. Technol. 44 (10), 3853–3860.
Leusch, F.D.L., et al., 2017. Analysis of the sensitivity of in vitro bioassays for androgenic, progestagenic, glucocorticoid, thyroid and estrogenic activity: Suitability for drinking and environmental waters. Environ. Int. 99, 120–130.
Leusch, F.D.L., et al., 2018. Comparison of in vitro and in vivo bioassays to measure thyroid hormone disrupting activity in water extracts. Chemosphere 191, 868–875.
Loos, R., 2012. Analytical methods relevant to the European Commission's 2012 proposal on Priority Substances under the Water Framework Directive. In: Kase, R., Kunz, P., Hollert, H., Werner, I. (Eds.), Contribution on bioanalytical assays for steroidal estrogens. ISBN 978-92-79-26642-3. 2012, Office of the European Union.
Loos, R., et al., 2018a. Analytical methods for substances in the Watch List under the Water Framework Directive. Publications Office of the European Union.
Loos, R., et al., 2018b. In: Loos, R., et al. (Eds.), Review of the 1st Watch List under the Water Framework Directive and recommendations for the 2nd Watch List. Publication Office of the European Union.
Moscovici, L., et al., 2020. Yeast-Based Fluorescent Sensors for the Simultaneous Detection of Estrogenic and Androgenic Compounds, Coupled with High-Performance Thin Layer Chromatography. Biosensors: 10th Anniversary Feature Papers 10 (11), 169.
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.
Neale, P.A., et al., 2017a. Integrating chemical analysis and bioanalysis to evaluate the contribution of wastewater effluent on the micropollutant burden in small streams. Sci. Total Environ. 576, 785–795.
Neale, P.A., et al., 2017b. Development of a bioanalytical test battery for water quality monitoring: Fingerprinting identified micropollutants and their contribution to effects in surface water. Water. Res. 123, 734–750.
Neale, P.A., et al., 2018. Solid-phase extraction as sample preparation of water samples for cell-based and other in vitro bioassays. Environ. Sci.: Processes Impacts 20 (3), 493–504.
Rechsteiner, D., et al., 2020. Natural estrogens in surface waters of a catchment with intensive livestock farming in Switzerland. Environ. Sci.: Processes Impacts 22 (11), 2244–2255.
Richter, J., et al., 2016. An interlaboratory comparison on whole water samples. Accreditation Qual. Assur. 21 (2), 121–129.
Sathishkumar, P., et al., 2020. Occurrence, interactive effects and ecological risk of diclofenac in environmental compartments and biota - a review. Sci. Total Environ. 698, 134057.
Schönborn, A., et al., 2017. Der Östrogenisierung der Umwelt auf der Spur: Wirkprofile östrogen-aktiver Substanzen in kommunalem Abwasser. Aqua & Gas 5.
Schriks, M., et al., 2015. DEMEAU Project Deliverable: Selection criteria to select in vitro bioassays for implementation and use.
Serra, H., et al., 2018. Triclosan Lacks Anti-Estrogenic Effects in Zebrafish Cells but Modulates Estrogen Response in Zebrafish Embryos. Int. J. Mol. Sci. 19 (4), 1175.
Severin, I., et al., 2017. Use of bioassays to assess hazard of food contact material extracts: State of the art. Food Chem. Toxicol. 105, 429–447.
Simon, E., et al., 2015. Challenges in effect-directed analysis with a focus on biological samples. TrAC Trends Analyt. Chem. 67, 179–191.
Simon, E., et al., 2019. Solid-phase extraction of estrogens and herbicides from environmental waters for bioassay analysis—effects of sample volume on recoveries. Analyt. Bioanalyt. Chem. 411 (10), 2057–2069.
Söffker, M., Tyler, C.R., 2012. Endocrine disrupting chemicals and sexual behaviors in fish – a critical review on effects and possible consequences. Crit. Rev. Toxicol. 42 (8), 653–668.
Sonneveld, E., et al., 2005. Development of androgen- and estrogen-Responsive Bioassays, Members of a Panel of Human Cell Line-Based Highly Selective Steroid-Responsive Bioassays. Toxicol. Sci. 83 (1), 136–148.
Sonneveld, E., et al., 2006. Comparison of In Vitro and In Vivo Screening Models for Androgenic and Estrogenic Activities. Toxicol. Sci. 89 (1), 173–187.
Tavazzi, S.C., et al., 2016. Water Framework Directive, Watch List Method - Analysis of 17b-estradiol and estrone. Publications Office of the European Union.
Thomas, K.V., et al., 2001. Characterization of estrogenic compounds in water samples collected from United Kingdom estuaries. Environ. Toxicol. Chem. 20 (10), 2165–2170.
van der Linden, S.C., et al., 2008. Detection of multiple hormonal activities in wastewater effluents and surface water, using a panel of steroid receptor CALUX bioassays. Environ. Sci. Technol. 42 (15), 5814–5820.
van der Oost, R., et al., 2017a. 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.
van der Oost, R., et al., 2017b. SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality assessment: Part II-field feasibility survey. Environ. Toxicol. Chem. 36 (9), 2400–2416.
Vermeirssen, E.L.M., et al., 2008. Estrogens in Swiss Rivers and Effluents & #8211; Sampling Matters. CHIMIA Int. J. Chem. 62 (5), 389–394.
Walker, C.W., Watson, J.E., 2010. Adsorption of Estrogens on Laboratory Materials and Filters during Sample Preparation. J. Environ. Qual. 39 (2), 744–748.
Wernersson, A.-S., et al., 2015. The European technical report on aquatic effect-based monitoring tools under the water framework directive. Environ. Sci. Eur. 27 (1), 7.
Williams, R.J., et al., 2012. Comparing predicted against measured steroid estrogen concentrations and the associated risk in two United Kingdom river catchments. Environ. Toxicol. Chem. 31 (4), 892–898.
Yarahmadi, H., et al., 2018. Seasonal variations of steroid hormones released by wastewater treatment plants to river water and sediments: Distribution between particulate and dissolved phases. Sci. Total Environ. 635, 144–155.
Zacs, D., et al., 2016. Determination of steroidal oestrogens in tap water samples using solid-phase extraction on a molecularly imprinted polymer sorbent and quantification with gas chromatography-mass spectrometry (GC-MS). Environ. Monit. Assess. 188 (7), 433.