

SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality assessment; Part I: model design and effect-based trigger values

Ron van der Oost^{†*}, Giulia Sileno^{†,‡}, Maria Suárez-Muñoz[‡], Mai Thao Nguyen[§], Harrie Besselink[¶] & Bram Brouwer^{¶,#}

[†] Waternet Institute for the Urban Water Cycle, Department of Technology, Research and Engineering, Amsterdam, The Netherlands

[‡] University of Amsterdam, Institute for Biodiversity and Ecosystem Dynamics (IBED), Amsterdam, The Netherlands

[§] Waterproef Laboratory, Research & validation, Edam, The Netherlands

[¶] Biodetection Systems BV, Amsterdam, The Netherlands

[#] Free University (VU), Institute for Ecological Sciences, Amsterdam, The Netherlands

*corresponding author:

Dr. Ron van der Oost

Waternet Institute for the Urban Water Cycle

Department of Technology, Research and Engineering

PO Box 94370,

1090 GJ Amsterdam,

The Netherlands

Email: ron.van.der.oost@waternet.nl

Phone: +31-652534378

Published in Environmental Toxicology & Chemistry 36: 2385-2399 (2017)

ABSTRACT

It is virtually impossible to reliably assess water quality with target chemical analyses only. Therefore, a complementary effect-based risk assessment is proposed here by bioanalyses on mixtures of bioavailable micropollutants: the Smart Integrated Monitoring (SIMONI) strategy. The goal of this strategy is to obtain more reliable information on the water quality to select optimum measures for improvement. The SIMONI strategy is two-tiered. Tier 1 is a bioanalytical hazard identification of sites. A tier 2 ecological risk assessment is carried out only at a limited number of sites where increased hazards are detected in tier 1. Tier 2 will be customized, based upon tier 1 evaluation and additional knowledge of the aquatic system.

The present paper focuses the tier 1 bioanalytical hazard identification to distinguish 'hot spots' of chemical pollution. First, a selection was made of relevant and cost-effective bioanalytical endpoints to cover a wide spectrum of micropollutant modes of action. Specific endpoints may indicate which classes of chemicals might cause adverse effects. Second, effect-based trigger values (EBT) were derived for these bioassays to indicate potential ecological risks. Comparison of EBT with bioassay responses should discriminate sites exhibiting different chemical hazards. Third, a model was designed to estimate the overall risks for aquatic ecosystems. The associated follow-up for risk management is a 'Toxicity Traffic Light' (TTL) system: green, low hazard (no action required), orange, potential risk (further research needed) and red high risk (mitigation measures). Thanks to cost-effectiveness, flexibility and relevance, the SIMONI strategy has the potential to become the first bioanalytical tool to be applied in regular water quality monitoring programs.

Key words: bioanalytical monitoring, micropollutants, bioassays, aquatic hazard & risk assessment, effect-based trigger values

1. INTRODUCTION

1.1 *Regular water quality monitoring and its limitations*

Water quality monitoring should be a control mechanism to assess i) whether or not measures are needed to improve quality, ii) which measures should be applied and iii) how successful these measures are in improving the water quality. The European Water Framework Directive (WFD) tries to integrate biological and chemical information in order to obtain an overall insight into the quality of individual water bodies. According to the WFD, the ecological status is determined by monitoring biological quality elements (absence and presence of taxa), hydro-morphological elements and physico-chemical elements of the water bodies. The chemical status of a water body is determined by analyzing the concentrations of 45 (groups of) priority substances. A good chemical status is reached when the concentrations of all substances are below the annual average and maximal allowable Environmental Quality Standards (AA- and MAC-EQs) defined to protect the environment and human health [1].

Regular chemical monitoring of the water quality is almost exclusively performed by targeted chemical analysis of a limited set of compounds. There are, however, some serious limitations related to the use of only chemical analyses of spot samples for monitoring the overall chemical status. First, because only a limited number of target substances are analyzed, the risk of non-target and unknown substances in the aquatic environment is unclear [2]. At present (December 2016), more than 125,000,000 substances are registered at the Chemical Abstracts Service (CAS), while already 20 years ago it was estimated that more than 100,000 chemicals are potential pollutants of aquatic ecosystems [3]. Second, it is obvious that chemicals do not occur as single substances in the environment, but in complex mixtures with potential synergism or antagonism. While concentrations of individual chemicals can be below lowest observed effect concentrations (LOEC) or detection limits, the entire mixture may still cause adverse effects [4]. Moreover, transformation products of micropollutants may be more toxic and persistent than the parent compounds [5]. These limitations may thus result in an incomplete assessment of the chemical hazards, e.g. [6], urging alternative approaches to be explored.

1.2 *Chemical quality monitoring with bioanalytical tools*

Effect-based monitoring tools to assess the chemical water quality by measuring effects instead of substances have already been applied for more than three decades.

Bioanalyses are being performed using two approaches, i) biomarkers in feral or caged organisms that are exposed in the field (e.g., reviews by Stegeman et al.[7] and Van der Oost et al., [6]) and ii) bioassays with laboratory organisms or cell-lines that are exposed to environmental samples or extracts (e.g., reviews by Castaño et al., [8] and Durand et al., [9]). Most in vivo assays (whole organisms) measure effects on gross parameters, such as growth, reproduction, feeding activity and mortality, while most in vitro assays (cell lines or unicellular organisms) measure specific biochemical effects of bioactive compounds, such as endocrine disruption and genotoxicity. Three major MoA groups can be distinguished regarding the type of interaction between a chemical and its molecular target: non-specific, specific and reactive toxicity. Non-specific toxicity or “baseline toxicity” refers to the minimum cytotoxicity that a chemical can exhibit not mediated by specific mechanisms (narcosis). Specific toxicity refers to all common mechanisms that involve the selective binding of a chemical to a protein (enzyme or receptor). MoAs are classified as reactive when covalent bonds are formed between the chemical and its target, or when chemical reactions such as oxidative stress are involved [10].

The added value of these effect-based tools for ecological risk assessment has been demonstrated in numerous studies (reviewed by Van der Oost et al., [6]). Firstly, in vivo assays respond to the presence of all pollutants in the water sample as well as their bioavailability and physical transfer into the test organism. Secondly, all bioanalytical tools give a more holistic assessment of biologically active chemicals present in the water, because they are able to detect mixture toxicity and the effects of metabolites and unknowns. For bioassays indicative of integrative effects (i.e. cytotoxicity, reactive and adaptive stress responses) typically only a very small fraction of the effect (often less than 1%) can be explained by known and identified chemicals [11, 12]. For other endpoints (e.g. estrogenic activity) it is easier to identify the compounds that cause the effects [13]. Water contaminants can elicit effects by interacting with critical cellular targets such as receptors, proteins, DNA or phospholipids that trigger a range of cellular events like the activation of genes, production of proteins, and altered protein signaling. Therefore, a series of pathways can be activated by contaminant exposure [14]. The concept of toxicity pathways is put into a wider ecotoxicological perspective as adverse outcome pathways (AOPs), linking the toxicity pathway at the cellular level, via responses at organ and organism level, ultimately to the response at the population level [15].

1.3 Requirements for a bioanalytical hazard identification

There are EU regulations that allow for bioanalytical methods in screening of feed and food for dioxin-like chemicals [16]. It is conceivable to adopt the bioanalytical

equivalency concept (BEQ) also for the development of bioanalytical trigger values for water quality assessment [11]. Yet, several conditions have to be met. A well-designed bioassay battery has the potential to provide a cost-effective assessment of the environmental risks caused by thousands of micropollutants. Several preconditions are important for the design of a monitoring strategy that is aimed to assess the potential ecological risks due to chemical pollution. The six criteria we considered most important for an effect-based monitoring strategy to be widely applied for regular water quality monitoring are:

- Identification of a broad spectrum of chemical pollutants. A hazard identification strategy based upon effect monitoring should be able to demonstrate the overall hazard of a wide range of chemical pollutants and their transformation products. The design of a good bioassay panel should thus cover the various types of toxic action, i.e. non-specific (various trophic levels), specific and reactive toxicity.
- Discrimination of sites with potential ecological health risks. The hazard assessment should use effect-based trigger values (EBT) to prioritize the sites where the highest ecological risks can be expected, but not all sites should be classified hazardous.
- Cost-effectiveness of the panel of bioassays in terms of equipment and consumables. Important for the acceptance of an alternative monitoring strategy is an attractive alternative that provides better ecological health-based information for the same budget or less.
- Good performance of bioassays, preferably with ISO or related validation level. The selected toxicological endpoints have to be measured in bioassays that meet certain performance quality standards, such as selectivity, accuracy, reproducibility, robustness, sensitivity, speed, and potential high-throughput capacity.
- Easy implementation and applicability of bioanalytical techniques by routine labs. The bioassays should be able to analyze environmental samples (complex mixtures), without high-tech laboratory requirements or specialist knowledge.
- Relevant and effective water sampling. Snapshot grab sampling is unreliable when water concentrations of micropollutants are varying. An alternative method is time-integrated sampling with passive samplers that are able to concentrate bioavailable micropollutants on site, and may be a good reflection of the micropollutants that accumulate in tissues of aquatic organisms [17, 18]. There are, however, certain pitfalls when combining passive sampling and bioassays [19] that will be discussed in a follow-up SIMONI paper [20]. Although online bioassays for real-time effects-based monitoring would be ideal for this purpose, current technologies are not sensitive enough to analyse the risks of low concentrations of micropollutants.

1.4 Objectives of the study

Due to the shortcomings of regular chemical monitoring there is an urgent need for more holistic evaluations of chemical water quality. In addition, monitoring should provide a more concrete quantitative risk classification, instead of the one-out-all-out principle of chemical EQS. The aim of the present paper is develop a strategy using complementary toxicological and chemical approaches, providing holistic risk quantification for chemical micropollutants to the aquatic environment. This strategy is not only appropriate for research projects, but can be applied in regular water quality monitoring. The present paper describes the design of the SIMONI (Smart Integrated Monitoring) strategy that meets the requirements outlined in section 1.3. This alternative for regular chemical monitoring does not advocate revolutionary new bioanalytical tools, but aims to find the optimal cost-effective combination of validated in vivo and in vitro endpoints to estimate environmental hazards by applying effect-based trigger values (EBT).

The primary objectives of the present paper were to describe the selection of relevant toxicological endpoints and the derivation of effect-based trigger values (EBT) for a suite of in vivo and in vitro bioassays, to be used in the Smart Integrated Monitoring (SIMONI) strategy for water quality assessment. This strategy is based upon a bioanalytical monitoring battery that is suitable to identify a wide range of chemical hazards. EBT for environmental risks were derived and used to interpret and classify the observed bioassay responses. In addition, a simple model is described that integrates all bioassay responses into a quantitative SIMONI score for hot spot assessment of ecological risks. The SIMONI strategy for environmental hazard assessment will allow regulators to link the bioassay results to potential adverse effects on the aquatic ecosystem. The strategy aims to establish a framework for the most cost-effective bioassay panel that reliably indicates the broad-spectrum chemical hazards for both invertebrate and vertebrate aquatic organisms.

2. SIMONI DESIGN AND BIOANALYTICAL ENDPOINTS

2.1 Design of the SIMONI strategy

A two-tiered strategy was designed, based upon the combination of field-exposed passive samplers and laboratory bioassay measurements (Figure 1). The first tier is hazard identification and the second tier is risk assessment. The strategy can also be applied for bioanalytical examination of concentrated large-volume water samples, in order to more accurately quantify the results (e.g. for tier 2 risk assessment).

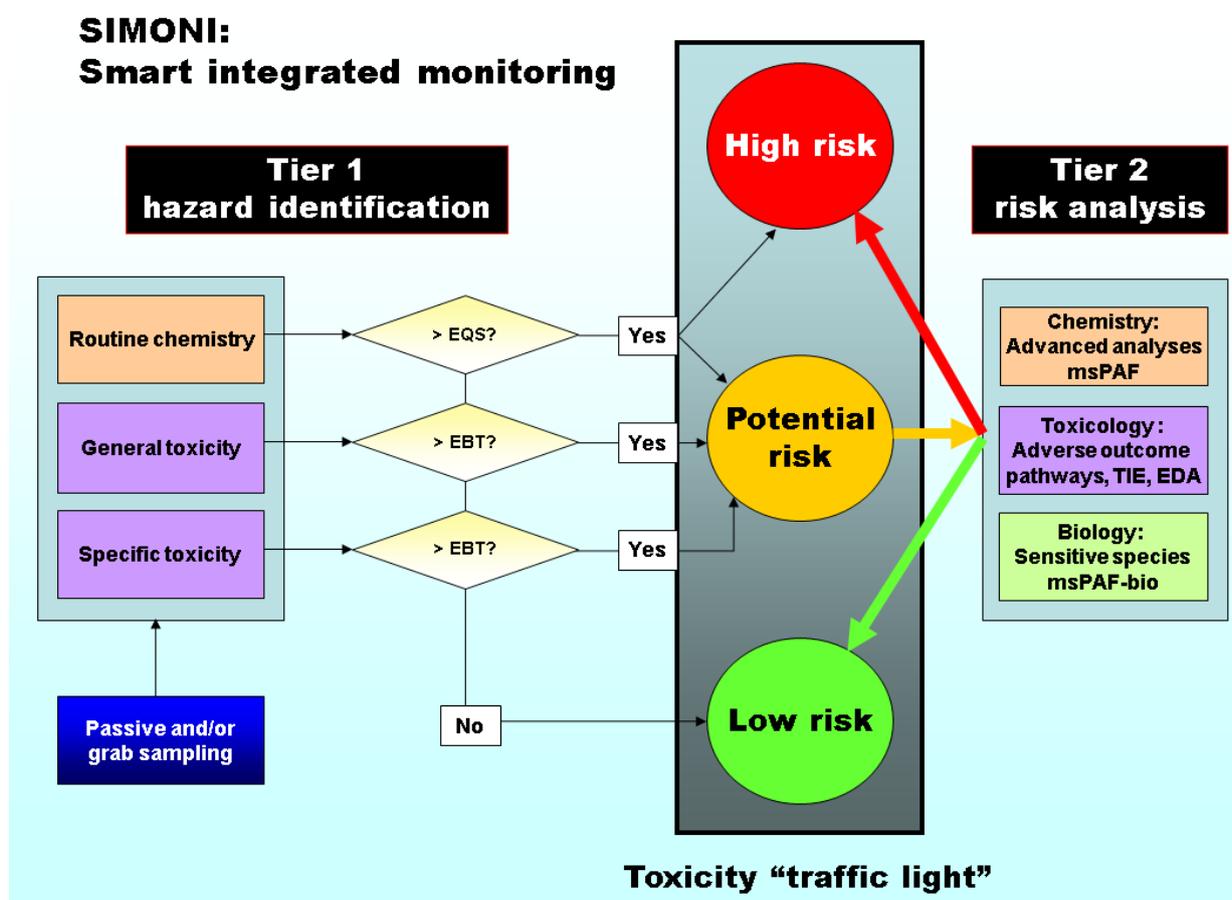


Figure 1: Schematic presentation of the SIMONI (Smart Integrated Monitoring) effect-based monitoring strategy; EQS = environmental quality standard; EBT = effect-based trigger value; msPAF = multiple substances potentially affected fraction of species; TIE = toxicity identification & evaluation; EDA = effect directed analysis

The first tier of the strategy, the hazard identification, is applied to assess the potential risks of a broad spectrum mixture of chemical micropollutants. The main objective of this screening phase is to identify the 'hot spots' of chemical water pollution. Hazards of organic micropollutants are characterized by evaluating the responses of a

suite of validated bioassays, using effect-based trigger values (EBT). In this tier, chemical analyses are only performed on metals and ammonium. Concentrations of inorganic compounds can be directly compared to EQS values for ecological risk assessment. A limitation of the strategy described in the present paper is that mixture effects determined with bioassays do not account for potential mixture interaction between organic and inorganic substances, except for a field *Daphnia* test that is exposed to surface water. Only a limited number of sites, where bioassay responses exceed EBT and indicate potential ecological risks, should be examined by tier 2, a more expensive second phase, for the actual risk assessment. All EBT exceedances indicate environmental hazards. However, if only a slight EBT exceedance is observed in a bioassay, this does not necessarily indicate an increased ecological risk. Therefore, a model is designed (section 5) which adds up EBT excesses in all bioassays and calculates an overall SIMONI score that should be indicative for the ecological risks. The tier 2 risk assessment phase should be customized, based upon tier 1 data and knowledge of the water system. If the tier 1 classification indicates low chemical hazards, there is no need to apply more advanced and expensive chemical analyses (e.g. WFD priority pollutants) and bioanalytical methods (e.g. fish bioassays and biomarkers). These analyses have to be carried out at sites with increased chemical hazards, where they are most relevant. If tier 2 chemical analyses do not give the answers needed to explain the effects observed in tier 1, a chemical-toxicological effect-directed analysis (EDA) can be performed using (in vitro) bioassays with the most pronounced responses to identify the unknown compounds that cause the observed toxicity. Results of the risk assessment can be verified with ecological observations, such as the occurrence of species that are sensitive to certain micropollutants.

2.2 Selection of bioanalytical endpoints

The selection of the bioanalytical endpoints for the SIMONI effect-based strategy was based upon requirements listed in section 1.3. For the detection of toxic responses of a broad spectrum of micropollutants it is important to select some non-specific in vivo assays, covering different trophic levels of aquatic organisms. In addition, the most relevant specific and reactive in vitro endpoints for water quality assessment have to be selected. A major advantage of specific in vitro responses is that these may indicate which classes of chemicals, and which types of effects may cause the main problem for aquatic organisms. The final selection of toxic endpoints was based upon literature, e.g. [21, 22, 23, 24], as well as our own research [20].

Willemsen et al. [21] evaluated the use of approximately 30 aquatic bioassays for non-specific toxicity, teratogenicity and genotoxicity. Selection criteria used for the

composition of a test battery were: 1. acute tests with whole organisms, 2. a small test volume; 3. tests available in kits; and 4. no specifically trained personnel or extensive laboratory facilities needed. Various methods were compared in terms of documentation, reproducibility, sensitivity, exposure time, standardization, technical simplicity and costs. A test battery was proposed using bacteria (Microtox®), algae (microplate assay), crustaceans (Thamnotoxkit F™) and a Daphnia (IQ) test.

Van der Linden et al. [22] investigated the possible adverse effects of endocrine disrupting compounds (EDC) with a panel of in vitro bioassays. The bioassays for estrogen (ER), androgen (AR) and glucocorticoid (GR) receptor activities were applied to extracts of municipal waste water treatment plant (WWTP) effluents. Different types of hormone receptor activity were detected and used to quantify the presence of EDC.

In a study by Macova et al. [23] a bioanalytical test panel was used for monitoring organic micropollutants across the complete water cycle from sewage to drinking water. Six endpoints targeting groups of chemicals with modes of toxic action relevant for human and environmental health were included in the evaluation: genotoxicity, endocrine disruption, neurotoxicity, phytotoxicity, xenobiotic metabolism and non-specific cell toxicity. All selected toxicity endpoints appeared to be relevant to evaluate the water cycle quality, and showed highest responses in WWTP influents. The effects in all six selected bioassays decreased across seven water treatment barriers.

In an extensive inter-laboratory study by Escher et al. [24] a representative set of water samples was analyzed for a broad range of toxicological effects by using 103 unique in vitro bioassays. Ten samples of WWTP effluent, recycled waters, storm water, surface and drinking waters were examined. Each water type had a characteristic bioanalytical profile. The most relevant modes of toxic action identified in that study were related to non-specific toxicity, xenobiotic metabolism (activation of aryl hydrocarbon [AhR] and pregnane X receptors [PXR]), hormone-mediated MoA (estrogenic, anti-androgenic and glucocorticoid activity) and reactive MoA (genotoxicity and oxidative stress).

Several bioassays were performed on surface water extracts of the Amsterdam region in The Netherlands [20]. Less than 50% of the polar extracts responded in the non-specific in vivo toxicity assays. Main in vitro responses of the polar extracts were estrogenic activity, anti-androgenic activity and genotoxicity. The non-polar extracts showed a higher percentage of detectable responses in the non-specific toxicity assays, especially in the Microtox and Daphniatox assays. Very high detectable response

percentages were obtained in the in vitro assays for dioxin- and PAH-like effects, oxidative stress (Nrf2), pregnane X response, estrogenic and anti-androgenic activity.

The selection of toxic endpoints for the SIMONI model is presented in Table 1. This endpoint selection aims to cover a broad range of micropollutants, multiple modes of action (non-specific, specific & reactive) and multiple biological levels (in vitro & in vivo). It is emphasized that this is a selection of bioanalytical endpoints, not bioassays. Bioassays that were used in the present study can be replaced by cost-effective alternatives that measure comparable endpoints and meet the criteria mentioned in section 1.3. Non-specific in vivo assays were included in the panel since they are responsive to the broadest range of micropollutants. Because of varying sensitivity for different organisms to different kinds of pollutants, assays with organisms at different trophic levels were chosen (bacteria [Microtox], algae [Algaetoxkit] and crustaceans [Daphniatoxkit]). In addition to the bioassay battery that was applied on concentrated water samples, a field *Daphnia magna* assay that assesses mortality after one week exposure to surface water was added to the non-specific endpoints. In vitro CALUX controls for cytotoxicity were also applied as apical endpoints. A fish embryo bioassay would fit very well into this panel, because it would be a whole-organism response measured in an in vitro assay. Due to the high costs, however, it did not meet the requirement of a cost-effective approach.

Table 1: Selection of SIMONI endpoints for effect-based hazard identification of micropollutants, with examples of targeted chemicals.

Category	Endpoint (mode of action)	Targeted chemicals	
Non-specific (in vivo)	Non-specific toxicity zooplankton, in situ	All chemicals	
	Non-specific toxicity zooplankton	All extracted chemicals	
	Non-specific toxicity phytoplankton	All extracted chemicals	
	Non-specific toxicity bacteria	All extracted chemicals	
	Non-specific toxicity cytotoxicity	All extracted chemicals	
	Specific (in vitro)	Estrogenic activity	Natural and synthetic estrogens, pseudo-estrogens, bisphenol A, alkyl phenols, pharmaceuticals, pesticides
		Anti-androgenic activity	Various pesticides, insecticides, herbicides, brominated flame retardants, (pseudo-) androgens, anabolic steroids, antibiotics, growth promoters, estrogens, PCBs
Glucocorticoid activity		Wide range of pharmaceuticals, corticosteroids	
Metabolism: Pregnane X receptor		Pesticides, PAHs, alkyl phenols, triazin pesticides, pharmaceuticals, polychloro biphenyls, cyanotoxins	
Metabolism: Aryl hydrocarbon receptor (persistent substances)		PCDDs, PCDFs, PCBs, brominated compounds	
Metabolism: Aryl hydrocarbon receptor (degradable substances)		PAHs, nitro-PAHs, halogenated PAHs	
Lipid metabolism (PPAR)		Organotins, perfluorinated compounds, esters, fatty acid derivatives, retinoic acid	
Antibiotic activity		Five classes of antibiotics (amidoglycosides, macrolides & β -lactams, sulfonamides, tetracyclines and quinolones), biocides (triclosan)	
Reactive (in vitro)	Genotoxicity	Chlorinated byproducts, aromatic amines, PAHs	
	Adaptive stress response: Oxidative stress	General chemical stress, reactive compounds, fungicides, insecticides, phenoles, pharmaceuticals, estrogens	

Specific in vitro responses were selected since these are generally much more sensitive than in vivo responses, and are able to detect specific activities caused by unknown mixtures of compounds with the same MoAs [25]. Endocrine disruptive effects were most frequently detected by estrogenic, anti-androgenic and glucocorticoid activities (e.g., ER, anti-AR and GR CALUX). Another promising endpoint, the (anti) progestagenic activity, was not included in the current selection, but will be evaluated in future research. Specific assays for activation of the aryl hydrocarbon receptor (AhR) responded to many degradable and persistent hydrocarbons (e.g., PAH and DR CALUX). Since lipid metabolism can be disrupted by environmentally relevant compounds, such as perfluorinated chemicals, peroxisome proliferation (e.g., PPAR γ CALUX) is considered a relevant endpoint; based on bioactivity profiling of environmental chemicals, the gamma isoform of the PPAR is considered the most relevant in environmental monitoring. Oxidative stress (e.g. Nrf2 CALUX) and activation of the pregnane X receptor (e.g. PXR CALUX) are relevant endpoints that respond to many WFD priority substances [26]. Finally, the activity of antibiotics (e.g. RIKILT WaterSCAN assay) is interesting to monitor, both for toxic responses to antibiotics and biocides and potential increase of resistant bacteria. Although CALUX bioassays were used for EBT derivation, it is important to mention that most selected in vitro endpoints can be tested by alternative technologies for specific toxicity measures (T47DKBluc, MDA-kb2, GeneBLAzer battery, MELN, MVLN, HG5LN battery, LUC battery, etc.). A disadvantage of many assays (including CALUX) is that license fees have to be paid for implementation. If relative effect potencies (REP) of key pollutants are significantly different in the alternative assays, EBT as derived in the present study have to be adjusted. This selection of SIMONI endpoints is intended to be updated as new insights emerge. New endpoints can be added if sufficient weight of evidence is available for detecting an important group of pollutants that is not covered by the initial selection. For implementation in the tier 1 screening, it should be possible to measure this endpoint with a bioassay that meets the criteria mentioned in section 1.3.

3. DEVELOPMENT OF EFFECT-BASED ENVIRONMENTAL TRIGGER VALUES

If bioanalytical tools are applied for water quality assessment it should be decided which bioassay response is considered to indicate an environmental hazard. To this purpose a suite of effect-based 'trigger values' (EBT) have been derived that differentiate between i) low risk for adverse ecological health effects if bioassay responses are below EBT, and ii) potential risk for adverse ecological health effects if bioassay responses exceed the EBT.

3.1 *Effect-based trigger values for non-specific toxicity*

The only bioassay that was applied on site in non-concentrated surface water was the in situ Daphnia assay. The percentage survival of 20 in situ exposed *Daphnia magna* was monitored after one week of exposure. An observed mortality of 20% was used as EBT for potential ecological effects, since this percentage is used as blank validity criterion for the chronic Daphnia assay [27].

An approach to derive environmental EBT for apical endpoints is described by Durand et al. [9]. These EBT were based on the assumption that acute toxicity in a concentrated sample provides an indication of chronic effects in the original sample. Although it is obvious that acute-to-chronic ratios (ACRs) may vary for different chemicals, the applied average ACR is 10, which is based on a comparison of the species sensitivity distributions using hundreds of tests for different species and substances [28, 29]. Chronic effects are estimated to occur at concentrations 10 times lower than the concentration where acute toxicity is observed (i.e. a detectable acute EC50 in a 10 times concentrated water sample), and negligible effects are expected at concentrations 100 times lower than the observed acute EC50 (Durand et al., 2009). Since most substances are not fully recovered by either SPE methods or passive sampling, the toxicity of the original sample may be underestimated in concentrated samples. EBT levels are therefore corrected for limited recovery by an assumed safety factor of 2 (50% recovery), as proposed by Durand et al. (2009). Taking this safety factor into account, effects measured below a relative enrichment factor (REF) of 20 are considered indicative of chronic effects, while REFs above 20 indicate a lower risk. If REF is converted to Toxic Units ($TU = 1/REF$), an EBT of 0.05 TU is proposed for potential chronic effects.

3.2 *Effect-based trigger values for specific and reactive toxicity*

Trigger values for in vitro bioassays should be derived by combining an approach based upon toxic equivalents (TEQs) or bioanalytical equivalents (BEQs) of selected substances that trigger the bioassays, together with a benchmark approach using known chemical, toxicological and biological data [30]. The rationale behind the benchmark approach is that bioassay responses observed at sites with a good ecological status should be considered as a background bioanalytical equivalent (BEQ) level of the bioassay. BEQ concentrations are a measure to express the effect of mixtures of unknown and potentially unidentified chemicals into the concentration of a known reference compound eliciting the same effect [2]. It is impossible to derive practically applicable trigger values that are safe for 100% of the aquatic organisms, as most of them would be exceeded even at unpolluted sites [20]. Therefore, a more realistic approach was applied in order to derive 'low-risk' effect-based trigger values. These specific EBT do not protect all aquatic organisms against adverse effects, but exceedances indicate elevated hazards for the aquatic ecosystem due to micropollutants.

The novel three-step approach for EBT derivation of all in vitro endpoints developed in the present study is schematically presented in Figure 2. In the following paragraphs the three steps will be explained one by one.

3.2.1 *Toxicological database with BEQ and acute-to-chronic conversions*

Since bioassays are effect-based tools that measure activities caused by a mixture of compounds, the identity of the compounds that cause the observed effect is unknown. Since using toxicological data for only one reference compound is therefore unreliable for most bioassay EBT derivations, a selection of other compounds that are able to trigger a response was made for all endpoints. Selection of compounds that have high relevance for EBT assessment was based upon toxicity, relative effect potency (REP) compared to reference compounds, available toxicity data and the range of reported concentrations in water. A complete list of REP values of the selected compounds for each bioassay is listed in Table S1 (Supplemental Data). Since chemicals with very low relative effect potencies (REPs) and high toxicity (low EC50) will give extremely low toxic BEQ values, a certain restriction was needed for realistic hazard identification:

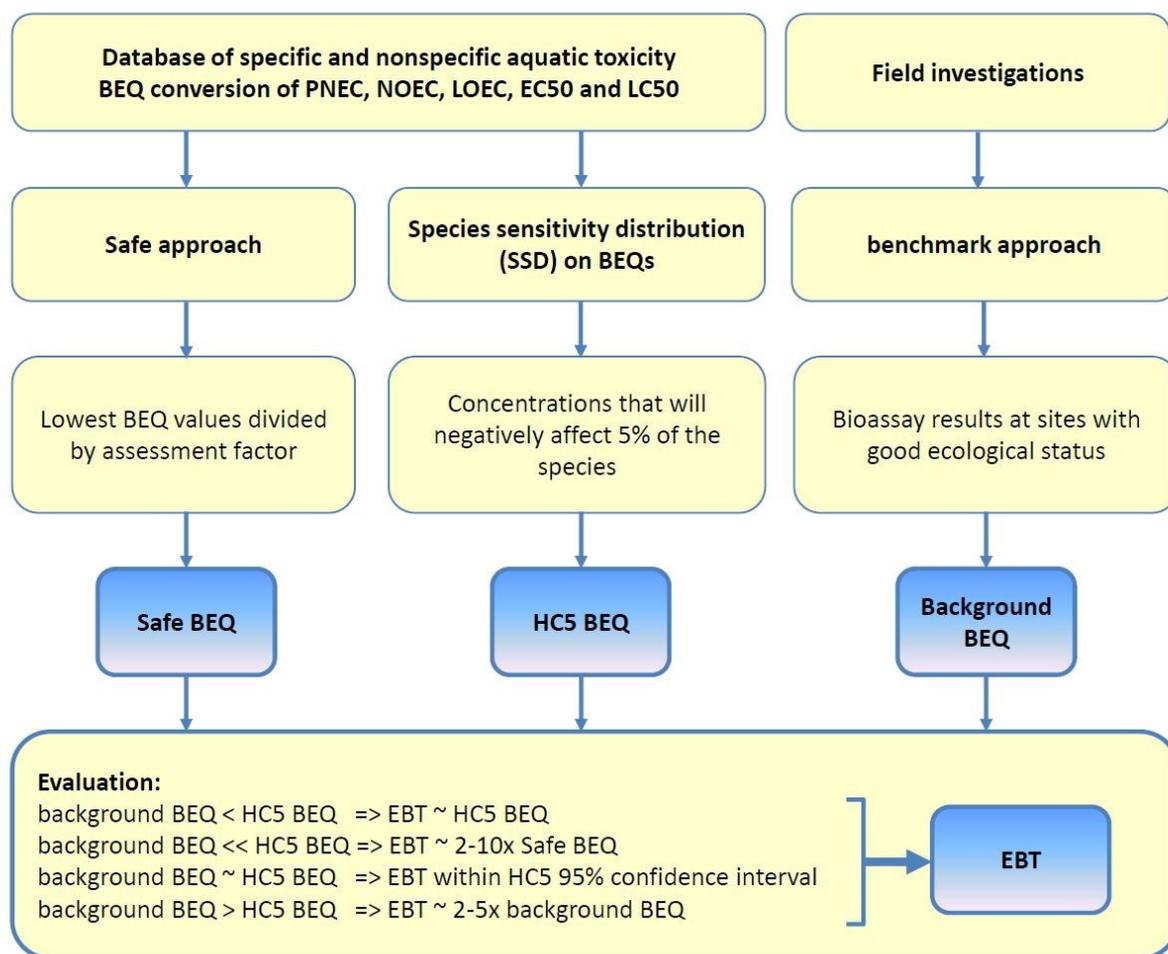


Figure 2: Schematic presentation of the approach for the derivation of effect-based trigger values (EBT); << means at least 100 times less

Assumption 1: in order to restrict the REP impact, all EBT procedures were performed on chemicals with REPs >0.001.

The REPs for the CALUX bioassays were calculated from PC10 results (concentration giving 10% induction with respect to the positive reference compound) from different compounds (provided by BioDetection Systems, Amsterdam, The Netherlands). The REPs for the antibiotic activities were estimated from the detection limits for the selected compounds in the RIKILT WaterScan (provided by RIKILT, Wageningen, Netherlands), as concentrations that cause a detectable bacterial growth inhibition should correspond to effect potencies.

Aquatic toxicity data (both specific and non-specific) were collected for all selected substances for a wide range of aquatic species. The toxicity databases were not restricted to population relevant endpoints (survival, growth and reproduction), or to specific

responses related to the mode of action of the bioassay. Toxic concentrations were converted to bioanalytical equivalents (BEQs) by multiplying them by the REP of the compound. Toxicity data, such as no observable effect concentration [NOEC], lowest observable effect concentration [LOEC], effect concentration causing 50% effect [EC50] or lethal concentration for 50% of the test animals [LC50], were used to build the datasets for each in vitro bioassay. Since predicted no effect concentrations [PNEC] are derived by using assessment factors, they cannot be directly compared to the other toxicity data. BEQ transformed PNECs, however, were also used to derive safe BEQ values. The complete dataset of all toxicity data that were used for the trigger value designs of all bioassays in the present study is presented in Table S2 (Supplemental Data). Using both acute and chronic toxicity data causes a discrepancy in data interpretation. Since chronic toxicity data were considered most relevant for environmental risk assessment, a second assumption was made in order to obtain a standardized acute-to-chronic data conversion.

Assumption 2: acute toxicity data were converted to chronic toxicity data by dividing them by an acute-to-chronic ratio of 10 [9]. Since there are no strict definitions for acute/chronic exposure times, it had to be decided to which category the obtained data were allocated, depending on the duration of the life cycles of the test species (estimations listed in Table 2).

Table 2: Criteria applied in the present study to estimate the duration of chronic exposure.

Organism	Chronic exposure (days)
Protozoans	
Bacteria	≥ 1
Fungi	
Polyps	
Algae	
Rotifers	
Crustaceans	≥ 4
Insects	
Molluscs	
Worms	
Plants	
Amphibians	≥ 7
Fish	

3.2.2 Safe BEQ assessment

Safe BEQs should indicate no-risk levels of active compounds to the ecosystem. The lowest chronic BEQ concentrations for each toxicological endpoint (NOEC, LOEC, EC50 and LC50) or the lowest PNEC were selected and divided by an assessment factor (AF), which ranged for 1 to 100 according to the toxic endpoint considered. Values for these AF or safety factors are proposed by the authors in consultation with other Dutch experts on aquatic toxicology (see acknowledgements).

Assumption 3: assessment factors to estimate safe BEQs by extrapolation of PNEC and four different toxicity parameters are listed in Table 3.

Table 3: Assessment factors (AFs) applied in the present study to convert toxicity data to the assumed safe BEQ levels.

Endpoint	AF
PNEC	1
NOEC	1
LOEC	5
EC50	10
LC50	100

The lowest of all observed chronic BEQ data, divided by its respective assessment factor, was considered as an ecologically 'safe BEQ' for the bioassay response. As an illustration of the safe approach, the collected toxicity BEQ data for compounds with a significant estrogenic activity (REP > 0.001) are presented in Figure 3. The safe BEQ determined for estrogenic activity is 0.007 ng EEQ/L, based upon the LOEC for vitellogenin induction in rainbow trout (*Oncorhynchus mykiss*) after chronic exposure to 3.3 ng estrone/L [31]. The Safe BEQ is derived after multiplication by the estrone REP of 0.01 and division by an AF of 5 for LOEC. The lowest BEQ-converted toxic concentrations found for the compounds with a significant response in each endpoint are listed in Table S3 (Supplemental Data). Graphic representations of all collected toxicity data used for the EBT design for all bioassays are presented in Appendix S4 (Supplemental Data).

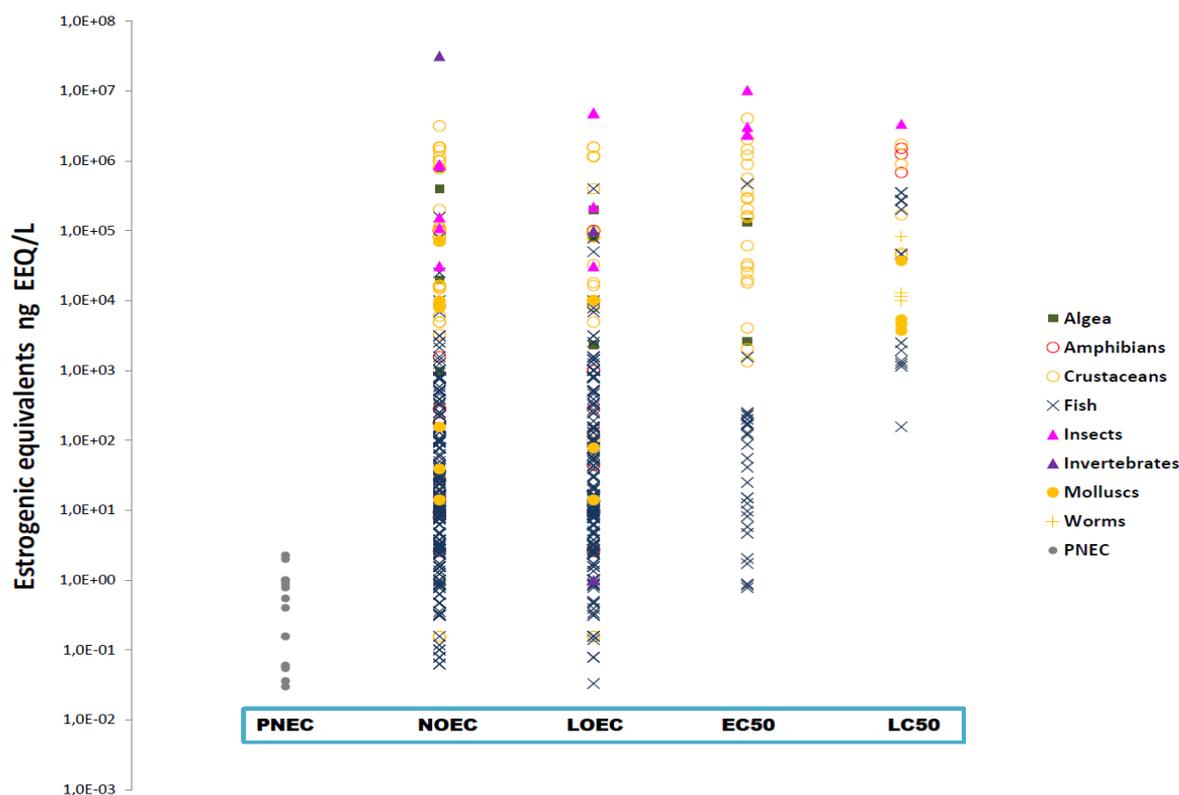


Figure 3: Bioanalytical equivalents (BEQ)-converted chronic toxicity data of compounds with estrogenic response

3.2.3 HC5 BEQ assessment

A more realistic trigger value approach ('low risk' instead of 'no risk') was based upon a Species Sensitivity Distribution (SSD) analysis [32], conform Escher et al. [33], who also applied a bioanalytical SSD analysis for recycled water EBT derivation. SSD curves are usually generated by fitting the distribution of average log-transformed toxicity data of several species for a single compound. The output of the SSD distribution can be used to determine the 5th percentile hazard concentration (HC5), which represents the concentration that will negatively affect 5% of the species. In the present study a standard SSD approach was hampered, because toxicity data of all compounds that trigger a bioassay response had to be included, and it is impossible to generate SSDs with data of different substances. To overcome this, all toxicity data were converted to BEQ concentrations of the reference compounds of the bioassay. This conversion allowed us to generate SSD curves with toxicity data for all species and all selected compounds. If different toxicity values were available for the same compound in the same species, the average BEQ values were used. The SSD curves were generated from EC50 BEQ values with the statistical software ETX 2.0 [34]. For DR CALUX and GR CALUX the number of available EC50 values was insufficient to run an SSD analysis, so additional NOEC and LOEC BEQs (multiplied by factors of 10 and 2, for conversion of NOEC and LOEC to EC50,

respectively) were used for these assays. The HC5 BEQ-values that were determined by this approach represent the intended upper limits of the low-risk effect-based trigger values (EBTs).

As an illustration of the SSD approach, the SSD curve with collected toxicity data (ng EEQ/L) for estrogenic compounds is presented in Figure 4. The BEQ-level that is hazardous to 5% of the organisms can be derived from the SSD curve (affected fraction 0.05). The HC5 BEQ for estrogenic compounds was estimated to be 0.5 ng EEQ/L. SSD curves of all collected toxicity data that were used for the EBT design for all bioassays are presented in Appendix S5 (Supplemental Data).

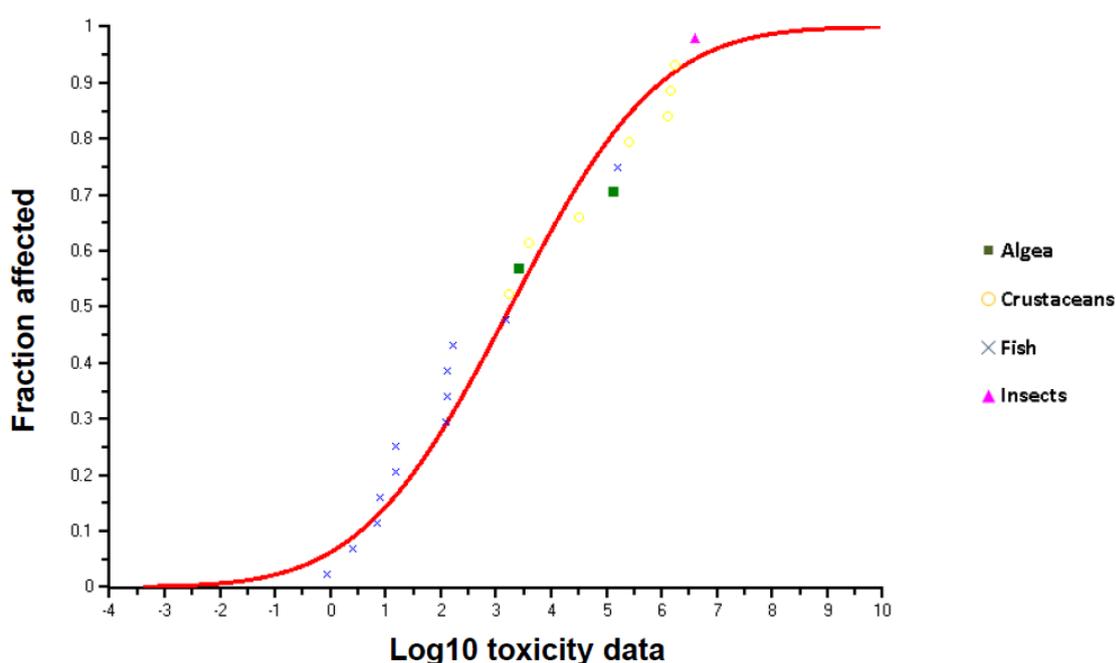


Figure 4: Species Sensitivity Distribution (SSD) of BEQ-converted toxicity data (ng EEQ/L of EC50) of estrogenic compounds; log10 toxicity data are expressed as ng EEQ/L.

3.2.4 Background BEQ assessment

Finally, a benchmark study with available field data was crucial in determining a realistic EBT. The rationale behind this approach was that average bioassay responses observed at sites with a good ecological status should be considered as a background level of the bioassay. The reasons for these background responses and the substances causing these effects at clean sites are unknown. Nevertheless, responses below these background BEQ levels are not considered indications for serious environmental hazards due to micropollutants, since no ecological problems were observed at the investigated healthy ecosystems. This approach was primarily based upon the results of a bioassay field survey

at eight Dutch reference sites with a good ecological status, according to WFD quality guidelines and/or expert opinion of Dutch ecologists. The background BEQs were determined as the average bioassay response at these ecologically healthy reference sites, as described in [20].

3.2.5 Evaluation of BEQ data and EBT derivation

The background BEQ data should ideally be in-between the safe BEQ and the HC5 BEQ. If the background BEQ was lower than the HC5 BEQ of the bioassay, than an EBT close to the HC5 BEQ was proposed. If the background BEQ was much lower than the HC5 BEQ of the bioassay, then a multiplication factor on the safe BEQ was proposed, depending on the strength of the dataset used to determine safe and HC5 BEQs. The multiplication factor varies between 2 (if toxicity data only for a limited number of substances and species were collected) and 10 (if sufficient information was available). If the background BEQ was close to the HC5 BEQ, then the proposed EBT was above the background BEQ but below the upper limit of the 95% confidence interval of HC5 BEQ. If the background BEQ was much higher than the HC5 BEQ, this could mean that the reference sites were (slightly) polluted or that the bioassays were triggered by confounding factors. Increased anti-AR responses and oxidative stress, for instance, were observed after exposure to natural organic matter (NOM) and humic acids of unpolluted soil [35, 36]. In these cases, trigger values based upon the lower HC5 BEQ would lose their discriminative power, since they are exceeded at all sites. Therefore, an EBT based on multiplication of the background BEQ (average of sites with a good ecological status) was proposed. The multiplication factor varies between 2 (if responses at polluted sites were close to the background BEQ) and 5 (if responses at polluted sites were much higher). This case was typically observed for bioassays that are responsive to a wide array of chemicals, such as anti-AR, oxidative stress and PXR responses. EBT derived for these assays are not considered indicative for low micropollutant risks, but are indicative for overall chemical stress.

3.3 Genotoxicity

Current guidelines for genotoxic substances that may be potentially carcinogenic assume that there is no safe level. The chances of adverse effects decrease at lower exposure levels, but a theoretical risk always remains. When interpreting bioassays for genotoxicity (e.g. Ames, umuC and p53 CALUX) the risks are assumed to be higher than for chronic toxicity. Since most genotoxicity bioassays are not easily quantifiable it was hard to use a BEQ approach for this endpoint. In consultation with other Dutch experts on aquatic toxicology (see acknowledgements), the EBT values as derived for the non-specific

bioassays (section 3.1) are further reduced by an assessment factor of 10. The proposed EBT for genotoxicity is 0.005 GTU (genotoxic units), i.e. a significant genotoxicity in a 200 times concentrated sample. This EBT was slightly exceeded at one of the eight clean sites (0.0065 TU); at other reference sites no genotoxicity was observed in 400x concentrated samples [20], supporting the choice made for 0.005 GTU.

4. RESULTS OF EBT DESIGN FOR NINE IN VITRO BIOASSAYS

The trigger value design of eight specific CALUX bioassays, as well as a RIKILT bioassay for five classes of antibiotics, will be described in this section. CALUX assays (Chemical Activated LUCiferase gene eXpression) are in vitro bioassays with genetically modified cell lines that contain specific receptor binding sites controlling expression of luciferase reporter genes. Antibiotic activities were analyzed by the RIKILT WaterSCAN bioassay using agar plates inoculated with five species of bacteria that are sensitive to five classes of antibiotics with different MoAs [37]. Both specific and non-specific effect concentrations were considered and toxicity data were not restricted to endpoints related to the mode of action of the bioassay. Since BEQ is the effect concentration times the relative effect potency (REP), it is therefore possible that the lowest converted BEQ (safe BEQ) is found for a low-REP compound with an entirely different MoA than the one typical for the bioassay. EBT were determined using an evaluation of safe, HC5 and background BEQ values, according to the algorithms described in section 3.2.5 and by expert judgement on varying multiplication factors.

4.1 *Estrogenic activity with ER α CALUX*

The toxicity of estrogenic compounds, including biomarker endpoints (e.g. production of vitellogenin and changes in gene expression) have been reported in many studies. Fish were found to be the most sensitive organisms to estrogenic effects. The safe BEQ was derived from a LOEC of 3.3 ng/L for vitellogenin induction in rainbow trout (*Oncorhynchus mykiss*) after chronic exposure to estrone [31]. After conversions, this yielded a chronic safe BEQ of 0.0066 ng EEQ/L (estradiol equivalents). The SSD analysis yielded a HC5 BEQ 0.5 ng EEQ/L. The average ER CALUX background BEQ was 0.06 ng EEQ/L. Since the background BEQ is lower than the HC5 BEQ, the proposed EBT for ER α CALUX activity is equal to the HC5 BEQ of 0.5 ng EEQ/L.

4.2 *Anti-androgenic activity with anti-AR CALUX*

The group of compounds that can inhibit the human androgen receptor and block its action (anti-androgenic response) is very heterogeneous, so many compounds were investigated (see Table S1, Supplemental Data). These included estrogenic compounds (e.g. 17 α -ethinylestradiol and estradiol), pesticides (e.g. alachlor, triclosan, vinclozolin),

synthetic materials (e.g. bisphenol A and phthalates) and non-ionic surfactants (alkylphenols). The safe BEQ was derived from the lowest LC50 of 0.016 µg/L after acute exposure of the copepod *Mesocyclops longisetus* to endosulfan [38]. After conversions, this yielded a chronic safe BEQ of 0.05 ng FluEQ/L (flutamine equivalents). The SSD analysis yielded a HC5 BEQ value of 0.13 µg FluEQ/L. The background BEQ at unpolluted sites was 4.55 µg FluEQ/L. Since the background BEQ is 35 times higher than the HC5 BEQ, the EBT derivation is based upon the background BEQ. Responses at polluted sites are clearly elevated, so the background BEQ is multiplied by a factor 5 to get an EBT for anti-AR CALUX activity of 25 µg FluEQ/L. This EBT is not indicative for low micropollutant risks, but can be considered as an indicator of overall chemical stress.

4.3 Glucocorticoid activity with GR CALUX

Since toxic effects of glucocorticoids on the aquatic community are poorly investigated, the toxicity dataset is limited. Most studies were conducted on fish, while information for other trophic levels is scarce or non-existent. The safe BEQ was derived from a LOEC for an increased gonadal somatic index in male fathead minnow (*Pimephales promelas*) after chronic exposure to 100 ng/L dexamethasone [39]. After conversions this yielded a chronic safe BEQ of 20 ng DexEQ/L. This safe BEQ is 3 orders of magnitude lower than the prednisolone PNEC of 27,800 ng DexEQ/L, derived with green algae toxicity and a safety factor of 1,000 [40]. The SSD analysis yielded a HC5 BEQ of 2145 ng DexEQ/L. The eight sites with good ecological status did not show any glucocorticoid activity above the detection limit of 1.2 ng DexEQ/L, so three orders of magnitude lower than the HC5 BEQ. Since the background BEQ is much lower than the HC5 BEQ, EBT derivation was based upon the safe BEQ. This safe BEQ was determined with a limited number of toxicity data, so a multiplication factor of 5 was used to get a proposed EBT for GR-CALUX activity of 100 ng DexEQ/L.

4.4 Dioxin-like activity with DR CALUX

Most dioxins and dioxin-like compounds are poorly water-soluble. They tend to accumulate in organisms due to bioaccumulation or biomagnification. Most of the studies reported nominal exposure concentrations, which may lead to an underestimation of the risk this group of compounds pose to aquatic organisms. Generally, samples for DR CALUX are cleaned-up by acid silica columns in order to remove easily biodegradable compounds such as PAHs. The safe BEQ was derived from the LOEC for reproduction of rare minnow (*Gobiocypris rarus*) after chronic exposure to 2 pg 2,3,7,8-TCDD/L [41]. After conversions,

this yielded a chronic safe BEQ of 0.4 pg TEQ/L (2,3,7,8-TCDD equivalents). The SSD analysis yielded a HC5 BEQ of 137 pg TEQ/L. The average background BEQ of the DR-CALUX response was 13 pg TEQ/L, which is 10 times lower than the HC5 BEQ. The EBT derivation is based upon the observed HC5 BEQ, but due to underestimations by nominal concentrations a lower BEQ was chosen for EBT derivation. The proposed EBT for overall dioxin-like activity is 50 pg TEQ/L, which is about three times lower than the HC5 BEQ.

4.5 PAH toxicity with PAH CALUX

As for dioxins, PAHs are lipophilic compounds that tend to accumulate in soil, organic particulate and tissues rather than dissolve in water. Therefore, the concentration of this class of pollutants should be measured during the exposure period, but the majority of studies reported nominal exposure concentrations. The PAH CALUX does not exclusively respond to PAHs, but also dioxins, PCBs and pharmaceuticals (e.g. cyclophosphamide) have significant REPs in the assay. With a limited exposure time of four hours, the PAH CALUX effect will mainly be caused by carcinogenic PAHs [42]. The safe BEQ was derived from a LOEC for reproduction of rare minnow (*Gobiocypris rarus*) after chronic exposure to 0.002 ng/L 2,3,7,8-TCDD [41]. After conversions, this yielded a chronic safe BEQ of 0.04 ng BaPenEQ/L (benzo[a]pyrene equivalents). The SSD analysis yielded a HC5 BEQ value of 41 ng BaPenEQ/L. The average background BEQ for the PAH CALUX was 63 ng BaPenEQ/L, which was close to the HC5 BEQ. The EBT derivation is largely based upon background BEQs, since PAHs are found everywhere, even at remote places, due to atmospheric deposition. The proposed EBT for overall PAH activity is 150 ng BaPenEQ/L, i.e. more than twice the background BEQ. This value is above the HC5 BEQ, but below the upper limit of its 95% confidence interval (Table 4).

4.6 Xenobiotic metabolism with PXR CALUX

The PXR CALUX is able to detect many WFD priority compounds, including pesticides, PAHs and alkyl phenols. The safe BEQ was derived from the LOEC for hemoglobin gene expression in *Daphnia magna* after acute exposure to 1 ng/L of the acetyl cholinesterase inhibitor chlorpyrifos-ethyl [43]. This value was converted to the chronic safe BEQ of 0.004 ng NicEQ/L (nicardipine equivalents). The SSD analysis yielded a HC5 BEQ of 8 ng NicEQ/L. The average background BEQ for PXR-CALUX biotransformation activity was 1.71 µg NicEQ/L, which is 200 times higher than the HC5 BEQ. EBT derivation was based upon the background BEQ. Since responses measured at polluted sites were close to the background BEQ, the proposed EBT for PXR CALUX is 3 µg NicEQ/L, which is approximately

twice the background BEQ. This EBT corresponds to 15 µg chlorpyrifos-ethyl equivalents per liter. As stated earlier, this EBT does not indicate low micropollutant risks and should be considered as an indicator of overall chemical stress.

4.7 Lipid metabolism with PPAR γ CALUX

The in vitro PPAR γ bioassays are able to detect compounds that activate the PPAR gamma receptor, including several classes of aquatic contaminants, such as organotins and perfluorinated compounds. The safe BEQ was based upon the lowest PNEC of 0.14 ng/L for the AhR inducer dibenzo[a,h]anthracene [44]. After conversion, this yielded a chronic safe BEQ of 0.00014 ng RosEQ/L (rosiglitazone equivalents). The SSD analysis yielded a HC5 BEQ of 0.3 ng RosEQ/L. The average background BEQ of PPAR γ CALUX was 4.37 ng RosEQ/L, which is an order of magnitude above the HC5 BEQ. EBT derivation was based upon the background BEQ. Since responses measured at polluted sites were close to the background BEQ, the proposed EBT for peroxisome proliferation is 10 ng RosEQ/L, which is approximately twice the background BEQ. The EBT was just above the upper limit of the HC5 BEQ 95% confidence interval (7 ng RosEQ/L, Table 4) and the highest response observed at the clean sites (9 ng RosEQ/L). This EBT corresponds to 12.5 µg tributyltin equivalents per liter. This EBT is not indicative for low micropollutant risks, but can be used to indicate overall chemical stress.

4.8 Oxidative stress with Nrf2 CALUX

It was not possible to find any aquatic toxicity data for the reference compound of the Nrf2 CALUX bioassay (i.e. curcumin), but sufficient information was available on the many other compounds that cause oxidative stress to cells and trigger a response in Nrf2 CALUX. The safe BEQ was derived from the lowest NOEC for decreased gene expression of gonadotropin-releasing hormone in transgenic Japanese Medaka (*Oryzias latipes*) after acute exposure to 0.001 µg/L estradiol [45]. After conversions this yielded a chronic safe BEQ of 0.006 ng CurEQ/L (curcumin equivalents). The SSD analysis yielded a HC5 BEQ of 0.034 µg CurEQ/L. The average background BEQ of the Nrf2 CALUX response for oxidative stress was 4.25 µg CurEQ/L. Because of the broad range of toxic concentrations of compounds that cause oxidative stress, the low HC5 BEQ value was exceeded at all reference sites. EBT derivation was based upon the background BEQ. Since responses measured at polluted sites were close to the background BEQ, the proposed EBT for overall oxidative stress activity is 10 µg CurEQ/L, which is approximately twice the background BEQ value. This EBT corresponds to 25 µg carbendazim equivalents per liter.

As for anti-AR and PXR endpoints, this EBT does not indicate low micropollutant risks and should be considered as an indicator for overall chemical stress. This mode of action is recommended to be included in any bioanalytical test battery [24].

4.9 Antibiotic activities with RIKILT WaterSCAN

According to their mode of action antibiotics are generally divided into five classes: amidoglycosides (A), macrolides & β -lactams (M+B), sulfonamides (S), tetracyclines (T) and quinolones (Q). The RIKILT WaterSCAN (SCreening for ANTibiotics) is a semi-quantitative bioassay, designed to determine activities of all five classes of antibiotics. Since the modes of action of these five antibiotics classes are different, separate EBTs were developed for each class. Due to their designed specific modes of antibacterial action, antibiotics are most hazardous to micro-organisms. REPs for the RIKILT assay could only be calculated for a limited number of antibiotic compounds.

4.9.1 Amidoglycosides

The safe BEQ of the A antibiotics was based on the lowest PNEC of 300 ng/L for neomycin [46]. No conversion to chronic safe BEQ in neomycin equivalents (NeoEQ) was needed. The SSD analysis resulted in a very high HC5 of 33 μ g NeoEQ/L. No detectable aminoglycosides activity (>90 ng NeoEQ/L) was found at the eight clean reference sites. EBT derivation was based upon the safe BEQ. Because of a limited toxicity dataset, the proposed EBT for amidoglycosides is 500 ng NeoEQ/L, which is approximately twice the safe BEQ.

4.9.2 Macrolides & β -lactams

The safe BEQ of the M+B antibiotics was based upon the lowest EC50 for growth inhibition of *Microcystis aeruginosa* after chronic exposure to 18 ng/L tiamulin [47]. After conversions, a chronic safe BEQ of 1.8 ng PenEQ/L (penicillin equivalents) was derived. The SSD analysis yielded a HC5 BEQ of 98 ng PenEQ/L. No detectable activities of macrolides and β -lactams (>1.4 ng PenEQ/L) were found at the eight clean reference sites. Large differences were observed between safe BEQ and HC5 BEQ values. EBT derivation was based upon the HC5 BEQ. The proposed EBT for M+B antibiotics is 50 ng PenEQ/L, i.e. half of the HC5 BEQ.

4.9.3 Sulfonamides

The safe BEQ for S antibiotics was based on the lowest LOEC found for growth inhibition of zebra fish embryos (*Danio rerio*) after acute exposure to 1000 ng/L sulfadiazine [48]. After conversions this yielded a chronic safe BEQ of 10 ng Suleq/L

(sulfamethoxazole equivalents). The SSD analysis yielded a very high HC5 BEQ of 67 µg Suleq/L. A sulfonamide activity of 37 ng Suleq/L was found at one of the reference sites, while no activity (>2 ng Suleq/L) was found at other clean sites. The background BEQ was 4.6 ng Suleq/L. EBT derivation was based upon the safe BEQ. Since a sufficient toxicity dataset was collected, the proposed EBT for sulfonamides is 100 ng Suleq/L, i.e. ten times the safe BEQ.

4.9.4 Tetracyclines

The safe BEQ of the T antibiotics was based upon the lowest PNEC of 170 ng/L derived for oxytetracycline [46]. This PNEC was based upon an acute LC50 of the algae *Selenastrum capricornutum* after exposure to 170 µg/L oxytetracycline [49], using a safety factor of 1000. No conversion to chronic safe BEQ in oxytetracycline equivalents (OxyEQ) was needed. The SSD analysis yielded a high HC5 BEQ of 27 µg OxyEQ/L. No detectable tetracycline activity (>22 ng OxyEQ/L) was found at the eight clean reference sites. EBT derivation was based upon the safe BEQ of a limited dataset. The proposed EBT for tetracyclines is 250 ng Suleq/L, which is approximately twice the safe BEQ.

4.9.5 Quinolones

Although it is not used as an antibiotic, toxicity data for triclosan (an antibacterial and antifungal substance) was also investigated, because it gives a clear response in the quinolones bioassay. The safe BEQ was based upon the lowest EC50 for growth inhibition of the algae *Pseudokirchneriella subcapitata* after acute exposure to 530 ng/L triclosan [50]. After conversions, this yielded a chronic safe BEQ of 0.53 ng FlqEQ/L (flumequine equivalents). The SSD analysis yielded an HC5 BEQ of 8.8 µg FlqEQ/L. No detectable quinolones activity (>44 ng FlqEQ/L) was found at any of the eight clean reference sites. Considering the high HC5 BEQ and the low safe BEQ due to triclosan, the proposed EBT for this group of antibiotics is 100 ng/L, which is two orders of magnitude above the safe BEQ. An exception of the algorithm that EBT should be around five times the safe BEQ was made here because the safe BEQ was only due to one compound (triclosan), while the proposed EBT is still much lower than HC5 BEQ.

4.10 Overall evaluation

The complete set of toxicity data used for the EBT design with literature references, lowest BEQ values, graphs with toxicity data and SSD-curves for all endpoints are presented in Appendices S2-5 (Supplemental Data). All relevant data for the EBT design, such as lowest observed BEQ values (safe BEQ), hazardous BEQ for 5% of the water organisms (HC5 BEQ with 95% confidence intervals), average background BEQ (observed

at sites with good ecological status), and the proposed EBT are summarized in Table 4. The extensive toxicity database that was created for the trigger value development can be applied to derive EBT for similar *in vitro* bioassays detecting the same modes of action, possibly after adjusting the relative effect potencies (REP) of the compounds. The derived EBTs of Table 4 can be applied for bioassays with similar endpoints if REP values are close to those of CALUX and RIKILT assays (Table S1, Supplemental Data).

Table 4: Derivation of SIMONI 1.2 environmental effect-based trigger values (EBT) for *in vitro* bioassays

Endpoint*	Safe BEQ endpoint/compound	HC5 BEQ ** (95% CI range)	Background BEQ	EBT
Estrogenic activity	0.0066	0.52	0.06	0.5
ERa CALUX [ng EEQ/L]	LOEC/estrone	(0.019-5.4)		
Anti-androgenic	0.00005	0.13	4.6	25
antiAR CALUX [µg FluEQ/L]	LC50/endosulfan	(0.05-0.27)		
Dioxin and dioxin-like	0.4	137	13.2	50
DR CALUX [pg TEQ/L]	LOEC/2,3,7,8-TCDD	(15-736)		
Glucocorticoid	20	2145	<LOD	100
GR CALUX [ng DexEQ/L]	LOEC/dexamethasone	(116-14311)		
PPARγ receptor	0.00014	0.3	4.4	10
PPARγ CALUX[ng RosEQ/L]	PNEC/dibenzo[a,h]anthracene	(0.002-6.9)		
Toxic PAHs	0.04	41	63	150
PAH CALUX [ng BaPEQ/L]	LOEC/2,3,7,8-TCDD	(2.5-254)		
Oxidative stress	0.000006	0.034	4.3	10
Nrf2 CALUX [µg CurEQ/L]	NOEC/estradiol	(0.008-0.11)		
Pregnane X receptor	0.000004	0.008	1.5	3
PXR CALUX [µg NicEQ/L]	LOEC/chlorpyrifos-ethyl	(0.002-0.024)		
Antibiotics activities:				
- Aminoglycosides	300	33222	<LOD	500
RIKILT [ng NeoEQ/L]	PNEC/neomycin	(1546-219614)		
- Macrolides & β-lactams	1.8	98	<LOD	50
RIKILT [ng PenEQ/L]	EC50/tiamulin	(13-470)		
- Sulphonamides	10	67037	4.6	100
RIKILT [ng SulEQ/L]	LOEC/sulfadiazine	(24675-148222)		
- Tetracyclines	170	27275	<LOD	250
RIKILT [ng OxyEQ/L]	PNEC/oxytetracycline	(8292-68544)		
- Quinolones	0.53	8759	<LOD	100
RIKILT [ng FlqEQ/L]	EC50/triclosan	(2197-26050)		
		Grey: EBT > HC5 BEQ	No toxic relevance, Indication for chemical stress	

*: expressed as equivalents of the reference compounds: EEQ = estradiol; FluEQ = flutamide; TEQ = 2378-TCDD; DexEQ = dexamethasone; RosEQ = rosiglitazone; BaPEQ = benzo[a]pyrene; CurEQ = curcumin; NicEQ = nicardipine; NeoEQ = neomycine; PenEQ = penicillin; SulEQ = sulfamethoxazole; OxyEQ = oxytetracyclin; FlqEQ = flumequine.

** : 95% confidence intervals (in parenthesis)

<LOD = all below limit of detection

5. SIMONI MODEL FOR OVERALL MICROPOLLUTANT RISKS

A final step in the SIMONI hazard assessment strategy was the development of a model that translates all bioassay responses into potential ecological hazards. This model, SIMONI 1.2, aims to derive an indication of the environmental risks of micropollutants based upon the results of individual bioassays, as well as for the total battery of bioassays (earlier versions 1.0 and 1.1 were not published). In order to obtain a broad and sensitive identification of the chemical risks, the model is based upon a panel of in vivo and in vitro bioassays that are performed on concentrated water extracts. The ecological hazards can be predicted when a substantial subset (or all) of the bioassays shown in Table 1 are applied. It is important to realize that the reliability of the outcome will be highest if the entire bioassay battery is applied.

A 'no risk' approach would require the use of the safe BEQ values for risk assessment. A more realistic approach for the tier 1 hazard identification is to apply the proposed 'low risk' EBT. If only slight exceedances of the EBT are observed in one or two bioassays, this does not necessarily indicate an increased ecological risk. However, with slight EBT excesses in many bioassays or large EBT exceedances, the hazard indications become stronger. Therefore, the model uses a simple formula that aims to quantify the combined ecological hazards due to micropollutants, by integrating all individual bioassay responses. All bioassays have been given a weight factor, i.e. 2 for apical toxicity endpoints (responsive to all micropollutants) and 1 for specific and reactive toxicity endpoints (responsive to specific groups of micropollutants). These weight factors were introduced to get a better balance between the impact of in vivo and in vitro bioassays. The total weight of the SIMONI bioassay selection is 20, i.e. 10 for in vivo and 10 for in vitro bioassays. If the endpoint selection is changed in the future, weight factors have to be adjusted in order to restore the in vivo vs. in vitro balance. As a requirement for a reliable result it is assumed that the total weight of the applied bioassays must be at least 10, and both in vivo and in vitro endpoints should be used. The SIMONI model divides all bioassay responses (TU or BEQ) by their associated EBT and multiplies them with this weight factor. Average relative responses are calculated for the five antibiotics assays and for the genotoxicity test with and without metabolic activation. Results are summed for all applied bioassays and divided by a percentage of the total weight of all applied bioassays.

$$\text{SIMONI score} = \frac{\sum_{i=1}^n \left(\frac{\text{bioassay response}_i}{\text{EBT}_i} \right) \times \text{weight}_i}{0.5 \times \text{total bioassay weight}}$$

A total SIMONI score above 1 indicates a potential ecological risk due to elevated concentrations of micropollutants. In this SIMONI score model, it is assumed that an increased hazard for the ecosystem occurs when the responses of all bioassays are, on average, more than 50% of the proposed trigger values, i.e. the factor 0.5 in the equation. The provisional weight factors and EBT percentage threshold are proposed by the authors in consultation with other Dutch experts on aquatic toxicology (see acknowledgements). The feasibility of the model with these assumptions will be demonstrated in a follow-up paper [20].

6. DISCUSSION

The SIMONI strategy aims to estimate broad-spectrum chemical risks for invertebrate and vertebrate aquatic organisms, by combining validated *in vivo* and *in vitro* endpoints. Three aspects are of main importance for a successful implementation of the SIMONI system in regular monitoring programs. First, the choice of the most relevant bioanalytical endpoints, second, the design of acceptable bioanalytical trigger values for potential risks, and third, a comparison with regular monitoring strategies that are mainly based upon chemical analyses. The first two aspects will be discussed below; the third aspect is discussed in a second paper on field feasibility studies of the SIMONI strategy [20].

6.1 Selection of bioanalytical endpoints

For a complete assessment of toxic hazards it is relevant to select a panel of *in vivo* and *in vitro* bioassays representing all MoA groups. The suggested SIMONI endpoints (Table 1) combine *in vivo* assays that are most relevant for invertebrate organisms, and *in vitro* bioassays with defined adverse outcome pathways that are most relevant for vertebrate organisms. The selected *in vivo* test organisms are ecologically relevant; bacteria are important for nutrient cycling, algae are primary producers and the invertebrates are primary consumers of algae, and food for small fish. Selected *in vitro* endpoints, such as effects on endocrine activities are known indicators for impaired reproduction (fecundity and spawning behaviour) and AhR induction and oxidative stress are related to tissue damage and fish mortality. The initial selection may be adjusted in the future if new knowledge becomes available, new assays are developed or expensive bioassays are transformed into more cost-effective formats. The impact of other relevant toxicity endpoints, such as immunotoxicity and neurotoxicity that act by many different pathways, is hard to estimate with one simple bioassay [2]. Reproductive and developmental toxicity are two classical examples for which there is currently no comprehensive *in vitro* model [2]. In this respect, fish embryo tests (FETs), would be a very relevant bioassay for higher trophic level organisms, but current costs of these assays are too high. The FET will, however, be implemented in the SIMONI tier 2 risk assessment phase. Moreover, it is possible to use multiplex-PCR to measure fish embryo biomarkers on more than 40 gene expressions (J. Legradi, VU University Amsterdam, Netherlands, unpublished data). If a large part of the *in vitro* bioassay screening could be replaced by fish embryo biomarker responses, this would make the assay very interesting for the tier 1 assessment.

6.2 Effect-based trigger values

Routine monitoring with bioassays is still hampered by the lack of reliable interpretation guidelines. The challenge is therefore to provide effect-based trigger values (EBT) that allow regulators to link the test results to possible adverse effects on environmental or human health [2]. The fact that in many studies only a small percentage of the effect observed in a bioassay can be explained by known (chemically determined) substances makes it imperative to derive such EBT [11, 12]. The present paper describes a format to derive trigger values for bioassays to distinguish between low and increased ecological risks due to organic micropollutants. We realize, however, that it is impossible to make a solid distinction between a 'good' and 'bad' chemical status if the identity of compounds causing the bioassay responses is unknown. Therefore, chemical identification on the tier 1 hot spots is performed in tier 2. The discriminative power of the first tier bioanalytical screening should be high enough to indicate potential hazards that have to be verified in a second tier (Figure 1).

Our intention to derive EBTs indicative for low toxic risks between the safe BEQ and the HC5 BEQ is best met for those bioassays that are triggered by a limited amount of specific substances. Increased estrogenic activity, for instance, is mainly due to exposure to four substances, estrone, 17 β -estradiol, estriol and 17 α -ethinylestradiol [13]. It is hard to derive toxicity based EBTs for the more 'promiscuous' endpoints, such as anti-androgenic activity, pregnane X metabolism and oxidative stress, due to the immense number of substances that are able to trigger these bioassays. These endpoints are, however, relevant to be taken up in the SIMONI panel as sensitive indicators of the overall chemical or toxic pressure. EBT for these bioassays were mainly based upon the benchmark approach using bioassay responses at sites with good ecological status.

A similar bioanalytical strategy as we propose for water quality is already in use for determination of food quality in Europe. Bioassays are being used for high-throughput screening of large amounts of food and feed samples. Only samples with bioassay responses exceeding effect-based action levels (BEQ trigger values) need to be confirmed by chemical analysis. A decision limit was derived for bioanalyses of dioxins, based on a GC/MS confirmation and the condition that the chance of false negatives should be less than 5%. This bioanalytical procedure is laid down in EU legislation [51]. Also for hormones and antibiotics established methods are available to regularly apply bioassays in food quality control, e.g., [52, 53].

A limited number of effect-based trigger values for water quality assessment can be found in literature. Most of these, however, are derived for human health risk of drinking water, and cannot be compared to the EBT developed in the present study. Brand et al. [54] derived some human health EBT for hormonal activity in drinking water. These EBT were derived for in vitro CALUX assays on estrogenic, androgenic, progestagenic and glucocorticoid activities. An EBT approach by Tang et al. [12] for cytotoxicity measured by the bioluminescence inhibition in *Vibrio fischeri* was based on combined effects of mixtures of regulated chemicals, according to the concentration addition model and a QSAR approach for non-specific baseline toxicity. Escher et al. [11] used a similar strategy to derive an EBT for the oxidative stress response pathway with the AREc32 cell line. Escher et al. [33] proposed a statistical EBT design with a read across from existing water quality guidelines. This EBT derivation was based upon matching the effect concentrations with chemical guidelines and filtering out appropriate chemicals that are responsive in a given bioassay at concentrations near the guideline values. Statistical methods were used to derive specific effect-based trigger bioanalytical equivalent concentrations (EBT-BEQ) for 11 bioassays that target receptor-mediated toxicity.

Most comparable to EBT of the present study are the environmental thresholds determined by Jarosova et al. [55], Kunz et al. [56] and Leusch et al. [57]. Jarosova et al. [55] derived estrogenic equivalents that are safe regarding major steroid estrogens (EEQ-SSE) in municipal WWTP effluents. This EBT was based on the assumption that only steroid estrogens are responsible for in vitro estrogenicity, using estrogenic REP, in vivo estrogenic PNECs and relative contributions to overall estrogenicity of WWTP effluents. Derived EEQ-SSEs for ER-CALUX varied from 0.2 to 0.4 ng EEQ/L for long-term exposure. The EU Commission proposed an AA-EQS of 0.4 ng/L for 17 β -estradiol, based on toxicity SSD for population relevant effects. Kunz et al. [56] proposed to use this AA-EQS as a trigger value for overall estrogenic activity. Leusch et al. [57] proposed threshold levels above which in vitro responses were expected to lead to adverse effects in exposed organisms. The threshold for estrogenic activity (0.07 ng EEQ/L) was based on the lowest PNEC*REP_{E-SCREEN} value of 4 estrogens (lowest value for estrone). If we applied this approach to the ERa CALUX REPs, this would lead to an EBT of 0.06 ng EEQ/L, which is almost ten times lower than the 'realistic' EBT of 0.5 ng EEQ/L we propose. The EBT derived in the present study was, however, close to the trigger values proposed by Jarosova and Kunz.

Leusch et al. [57] also proposed threshold levels for non-specific toxicity determined with Microtox (1 TU) and genotoxicity determined with umuC assay (1 GTU), which were less conservative than the ones proposed in the present study (0.05 TU and 0.005 GTU,

respectively). These two trigger values of Leusch et al. were, however, based upon acute toxic effects while those derived in the present study were based upon chronic toxicity estimations, with an additional safety factor for genotoxic effects.

Hamers et al. [58] obtained toxicity pathway profiles as toxicological 'fingerprints', using a bioassay battery with different modes of action (genotoxicity, [anti]-estrogenic activity, thyroid activity, dioxin-like activity and non-specific cytotoxicity). Three potential approaches were described: i) toxicity profiles were translated into hazard profiles, indicating the relative distance to the desired quality status for each toxic mode of action, ii) toxicity profiles were translated into ecological risk profiles, i.e. the ratio between the bioassay responses and those considered safe for environmental health, and iii) use of toxicity and hazard profiles to select hot spot samples for further in-depth effect-directed analysis (EDA). A combination of the second and third approach is most similar to the SIMONI strategy proposed in the present paper.

6.3 Limitations of the bioanalytical approach

Both over- and underestimations of the overall toxic impact of the mixture can be made with the BEQ concept. Moreover, it is hard or impossible to predict adverse in vivo effects with in vitro responses, if the impact of absorption, distribution, metabolism and excretion (ADME) is unknown. It must be stressed that an effect-based trigger value should not be used as stand-alone, but a battery of biological endpoints should be combined.

There are several assumptions and limitations connected to the EBT design of the present study that may be optimized in future research. Three assumptions had to be made in the EBT design, i.e. acute-to-chronic toxicity ratio and duration of chronic studies, assessment factors to convert toxicity data to safe values, and the REP-based selection of compounds. Limitations with regard to the toxicological database are that only chemicals with a known REP in the bioassay could be considered for the BEQ approach. Reliable REP values are essential for a good conversion of substance concentrations to BEQ, but it was not always possible to assess REPs with EC50 data due to cytotoxicity in the bioassay at higher concentrations. The CALUX REPs were estimated with PC10 values (10% of the maximum bioassay response of the reference compound). Future research should be aimed at refining bioassay REP values of the most relevant compounds. The publically available REP database from the Toxicity ForeCaster (ToxCast) project on the US EPA website could be very useful on this matter. An additional limitation is that there may be large deviations between in vivo and in vitro REP values, e.g. [59].

For some of the bioanalytical endpoints background BEQs are higher than HC5 BEQ, which resulted in EBT that are not indicative for low micropollutants risks. Since it might be possible that the background BEQs are different in other regions, it is advised to perform additional background studies on local reference sites.

For several compounds it was hard or impossible to find data on aquatic toxicity, which resulted in limited datasets for some of the bioassays, e.g. GR CALUX and some antibiotics. All toxicity data used for EBT derivation were concentrations in the water that indicated a certain effect to aquatic organisms or safe PNEC values. Laboratory exposures with single compounds, however, differ from the situation in the natural environment due to varying physico-chemical conditions, such as pH, temperature and light exposure. In addition, the majority of the studies investigating the toxicity of highly hydrophobic compounds reported nominal concentrations, which may give underestimations of toxicity. This is, however, also the case when bioassays are applied on environmental extracts.

6.4 *SIMONI model for integrated risk assessment*

The holistic approach of effect-based toxicity monitoring has many advantages over the substance-based chemical monitoring of a limited number of target compounds. However, the effect-based methods are hardly used in regular monitoring programs, due to the lack of established guidelines for interpretation of the data. The integrated monitoring strategy presented in the present paper aims to provide these guidelines.

The in vivo and in vitro bioassay panel that is proposed for hazard identification in water concentrates, together with a field-exposed in situ Daphnia assay, should be able to distinguish chemical quality of sites indicating low or increased ecological hazards. Despite its limitations and uncertainties, the authors believe that the approach proposed in the present study constitutes a better alternative to the current EU WFD monitoring, i.e. chemical analyses of a limited amount of priority substances, while the potential impact of more than 100,000 unknown compounds remains unknown. The alternative strategy proposed in the present paper represents one of the first attempts to connect multiple bioassay responses with potential negative effects for aquatic organisms, by using effect-based trigger values (EBT) for low ecological risks.

In comparison to regular monitoring programs, the SIMONI strategy can be more cost-effective. A WFD chemical surveillance monitoring consists of 12 monthly grab samples that are analyzed for 45 priority substances. A WFD chemical campaign for one

water body currently (2016) costs around €40,000 in the Netherlands. The suggested battery of bioassays in Table 1 (total ~€2,000) is less expensive than chemical analysis of 45 WFD priority pollutants (total ~€3,000). The SIMONI strategy combines passive sampling with bioanalytical methods. Since several weeks of passive sampling will obtain time weight averages concentrations of bioavailable micropollutants, less sampling campaigns are assumed necessary for representative water quality assessment. When the entire panel of bioassays is analyzed, three seasonal SIMONI campaigns in one water body costs around €10,000. A more expensive second tier risk assessment, using effect-directed analyses (EDA, [60]), should only be performed at the sites with elevated ecological risks.

The SIMONI model estimates an integrated quantitative measure for potential ecological risks due to micropollutants. The outcome of the model indicates which sites are hot-spots, relevant for additional chemical-toxicological research. Moreover, if specific or reactive activities are above EBT, the model will indicate which class of chemicals may cause the main problem for aquatic organisms. The associated follow-up for risk management is the 'Toxicity Traffic Light' (TTL) system:

- Green, SIMONI-score <1: low hazard, so no action is required;
- Orange, SIMONI-score ≥1: potential risk, so tier 2 research should be performed;
- Red, SIMONI-score ≥1 + tier 2 identification toxic substances: high risk, so mitigation measures are needed; SIMONI-scores ≥2 (responses of all bioassays are, on average, more than 100% of the EBT) are also indicative for high risks, but most appropriate measures can be taken if the (class of) chemicals causing the effects is known.

If the information from the SIMONI hazard assessment is combined with other influences of other ecological key factors, a tailor-made plan can be designed for a tier 2 ecological risk assessment. The next steps to make the proposed concept attractive for risk assessors would be to gain experience upon the applicability to case studies and to evaluate its robustness for practical use. Field validation studies with this strategy are described in the second paper of the SIMONI series [20]. Due to its low costs and high relevance, this model has the potential to become the first bioanalytical strategy to be applied in regular monitoring of surface water quality. Most Dutch water authorities started feasibility studies with the SIMONI strategy in 2016.

7. SUPPLEMENTAL DATA

Table S1: Relative effect potencies (REP) of all compounds investigated for EBT development in all bioassays

TableS2: Toxicological data considered for EBT design in vitro bioassays

Table S3: Lowest toxicity values found for REP > 0.001 compounds

Appendix S4: Graphic representations of the toxicity data collected for each bioassay

Appendix S5: Graphs of SSD analyses of all in vitro bioassays

8. ACKNOWLEDGMENTS

The authors thank the Waternet Research & Innovation group for funding the main part of the project. Laura Moria (Waternet Institute for the Urban Water Cycle, Netherlands) is acknowledged for her useful input on practical applications of the strategy. We thank the STOWA Foundation for Applied Water Research for funding a study on EBT development and finalizing the SIMONI 1.2 modelling. We like to thank Dick de Zwart and Leo Posthuma (both National Institute for Public Health and the Environment, Netherlands), Jaap Postma (Ecofide Environmental and Ecological Consultancy, Netherlands), and Leonard Osté (Deltares, Dutch Institute for Applied Water Research, Netherlands) for consultancy on the assumptions for trigger value derivation. Tim Williams (University of Birmingham, UK), thanks a lot for critically reading and adjusting the manuscript. The authors especially want to thank Beate Escher (Centre for Environmental Research, University of Tübingen, Germany and University of Queensland, Australia) for her extensive review and useful comments that significantly improved the quality of the manuscript.

9. REFERENCES

- 1: European commission, 2012. Proposal for a directive of the european parliament and of the council amending Directives 2000/60/EC and 2008/105/EC as regards priority substances in the field of water policy. EC document COM/2011/0876 final: URL: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011PC0876&rid=3>
- 2: Escher BI and Leusch FDL, 2012. Bioanalytical Tools in Water Quality Assessment. IWA publishing, London (UK).
- 3: Depledge MH and Fossi MC, 1994. The role of biomarkers in environmental assessment (2). *Ecotoxicology* 3: 161-172.
- 4: Silva E, Rajapakse N, Kortenkamp A, 2002. Something from "nothing"-eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. *Environ Sci Technol.* 36:1751-1756.
- 5: Escher BI, and Fenner K, 2011. Recent advances in the environmental risk assessment of transformation products. *Environ. Sci. Technol.* 45: 3835-3847.
- 6: Van der Oost R, Beyer J, and Vermeulen NPE, 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. *Environ. Toxicol. Pharmacol.* 13: 57-149.
- 7: Stegeman JJ, Brouwer M, Di Giulio RT, Förlin L, Fowler BA, Sanders BM and Van Veld PA, 1992. Molecular responses to environmental contamination: enzyme and protein systems as indicators of chemical exposure and effect, in: *Biomarkers: Biochemical, physiological and histological markers of anthropogenic stress.* RJ Huggett, RA Kimerly, PM Mehrle Jr. and HL Bergman, Eds., Lewis Publishers Inc., Chelsea, MI, USA. pp. 235-335.
- 8: Castaño A, Bols, N, Braunbeck T, Dierickx P, Halder M, Isomaa B, Kawahara K, Lee LEJ, Mothersill C, Pärt P, Repetto G, Sintes JR, Rufli H, Smith R, Wood C and Segner H, 2003. The Use of Fish Cells in Ecotoxicology; the Report and Recommendations of ECVAM Workshop. *ATLA (Alternatives To Laboratory Animals)* 31: 317–351.
- 9: Durand AM, Rotteveel S, Collombon MT, Van der Grinten E, Maas JL and Verweij W, 2009. Toxicity measurements in concentrated water samples; evaluation and validation. RIVM (National Institute for Public Health and the Environment) report 607013010/2009.URL: http://www.rivm.nl/Documenten_en_publicaties/Wetenschappelijk/Rapporten/2009/mei/Toxicity_measurements_in_concentrated_water_samples_Evaluation_and_validation?sp=cm12bXE9ZmFsc2U7c2VhcmNoYmFzZT01OTQwMDtyaXZtcT1mYWxzZTs=&pagenr=5941.

- 10: Escher BI, and Hermens JLM, 2002. Modes of action in ecotoxicology: their role in body burdens, species sensitivity, QSARs, and mixture effects. *Environ. Sci. Technol.* 36: 4201-4217.
- 11: Escher BI, Van Daele C, Dutt M, Tang JYM, and Altenburger R, 2013. Most Oxidative Stress Response In Water Samples Comes From Unknown Chemicals: The Need For Effect-Based Water Quality Trigger Values. *Environ. Sci. Technol.* 47: 7002–7011.
- 12: Tang JYM, McCarty S, Glenn E, Neale PA, Warne MSJ, and Escher BI, 2013. Mixture effects of organic micropollutants present in water: Towards the development of effect-based water quality trigger values for baseline toxicity, *Water Research* 47: 3300-3314.
- 13: Aerni HR, Kobler B, Rutishauser BV, Wettstein FE, Fischer R, Giger W, 2004. Combined biological and chemical assessment of estrogenic activities in wastewater treatment plant effluents. *Anal. Bioanal. Chem.* 378: 688–696.
- 14: Collins FS, Gray GM, and Bucher JR, 2008. Transforming Environmental Health Protection. *Science (policy forum)* 319: 906-907.
- 15: Ankley GT, Bennet RS, Erickson RJ, Hoff DJ, Hornung MW, Johnson RD, Mount DR, Nickols, DW, Russom CL, Schmieder PK, Serrano JA, Tietge JE, and Villeneuve DL, 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* 29: 730–741.
- 16: Hoogenboom LAP, Haedrich J, Eppe G, Goeyens L, Elskens M, Malagocki P, Scippo M, Vanderperren H, Windal I, Kotz A, Denison M, Malisch R, 2010. Revised EU-criteria for applying bioanalytical methods for screening of feed and food for dioxins and dioxin-like PCBs. *Organohal. Compounds* 72: 1800-1905.
- 17: Smedes F, 2007. Monitoring of Chlorinated Biphenyls and Polycyclic Aromatic Hydrocarbons by Passive Sampling in Concert with Deployed Mussels. In: *Comprehensive Analytical Chemistry*, Eds. R Greenwood, G Mills and B Vrana, Elsevier, Volume 48, pp 407-448.
- 18: Li J-Y, Tang JYM, Jin L and Escher BI, 2013. Understanding bioavailability and toxicity of sediment-associated contaminants by combining passive sampling with in vitro bioassays in an urban river catchment. *Environm. Toxicol. Chem* 32: 2888-2896.
- 19: Jahnke A, Mayer P, Schäfer S, Witt G, Haase N and Escher BI, 2016. Strategies for Transferring Mixtures of Organic Contaminants from Aquatic Environments into Bioassays. *Environ. Sci. Technol.* 50: 5424-5431.

- 20: Van der Oost R, Sileno G, Janse T, Nguyen MT, Besselink H, and Brouwer A, 2017. SIMONI (Smart Integrated Monitoring) as a novel bioanalytical strategy for water quality assessment: Part II. Field feasibility survey. *Environ. Toxicol. Chem.* 36: 2400-2416.
- 21: Willemsen A, Vaal MA, De Zwart D, 1995. Microbiotests as tools for environmental monitoring. RIVM report 607042005.
- 22: Van der Linden S, Heringa M, Man H-Y, Sonneveld E, Puijker LM, Brouwer A and Van der Burg B, 2008. Detection of Multiple Hormonal Activities in Wastewater Effluents and Surface Water, Using a Panel of Steroid Receptor CALUX Bioassays. *Environ. Sci. Technol.* 42: 5814-5820.
- 23: Macova M, Toze S, Hodgers L, Mueller JF, Bartkow M, Escher BI, 2011. Bioanalytical tools for the evaluation of organic micropollutants during sewage treatment, water recycling and drinking water generation. *Water Res.* 45: 4238-4247.
- 24: Escher BI, Allinson M, Altenburger R, Bain PA, Balaguer P, Busch W, Crago J, Denslow NA, Dopp E, Hilscherova K, Humpage AR, Kumar A, Grimaldi M, Jayasinghe BS, Jarosova B, Jia A, Makarov S, Maruya KA, Medvedev A, Mehinto AC, Mendez JE, Poulsen A, Prochazka E, Richard J, Schifferli A, Schlenk D, Scholz S, Shiraishi F, Snyder S, Su G, Tang JYM, Van der Burg B, Van der Linden SC, Werner I, Westerheide SD, Wong CKC, Yang M, Yeung BHY, Zhang X and Leusch FDL, 2014. Benchmarking Organic Micropollutants in Wastewater, Recycled Water and Drinking Water with In Vitro Bioassays. *Environ. Sci. Technol.* 48: 1940-1956.
- 25: Sonneveld E, Jansen H J, Riteco JAC, Brouwer A, Van der Burg B, 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: 136-148.
- 26: Neale PA, Ait-Aissa S, Brack W, Creusot N, Denison MS, Deutschmann B, Hilscherová K, Hollert H, Krauss M, Novák J, Schulze T, Seiler T-B, Serra H, Shao Y and Escher BI, 2015. Linking in Vitro Effects and Detected Organic Micropollutants in Surface Water Using Mixture-Toxicity Modeling. *Environ. Sci. Technol.* 49: 14614-14624.
- 27: OECD, 2008. Test No. 211: *Daphnia magna* Reproduction Test, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789264070127-en>
- 28: De Zwart D, and Sterkenburg A, 2002. Toxicity-based assessment of water quality. Published in: *Species Sensitivity Distributions in Ecotoxicology*. Posthuma L, Suter II GW, Traas TP (eds.). Lewis Publishers, Boca Raton, FL, pp. 383-402.

- 29: Raimondo S, Montague BJ, and Barron MG, 2007. Determinants of variability in acute to chronic ratios for aquatic invertebrates and fish. *Environ. Toxicol. Chem.* 26: 2019-2023.
- 30: Maas JL, Van de Plassche E, Straetmans A, Vethaak AD, and Belfroid AC, 2003. Normstelling voor bioassays. Uitwerking voor oppervlaktewater en waterbodem. RIZA report 2003.005, (in Dutch).
- 31: Thorpe KL, Cummings RI, Hutchinson TH, Scholze M, Brighty G, Sumpter JP, Tyler CR, 2003. Relative potencies and combination effects of steroidal estrogens in fish. *Environ. Sci. Technol.* 37(6):1142-1149.
- 32: Posthuma L, GW Suter II, and Traas TP, 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, Boca Raton, FL. 587 pp.
- 33: Escher BI, Neale PA, and Leusch FDL, 2015. Effect-based trigger values for in vitro bioassays: Reading across from existing water quality guideline values. *Water Research* 81: 137-148.
- 34: Van Vlaardingen PLA, Traas TP, Wintersen AM, Aldenberg T (2004). ETX 2.0 – A Program to Calculate Hazardous Concentrations and Fraction Affected, Based on Normally Distributed Toxicity Data. RIVM (National Institute for Public Health and the Environment), report 601501028/2004
- 35: Bittner M, Saul N, and Steinberg CEW, 2012. Antiandrogenic activity of humic substances. *Sci. Total Environ.* 432: 93-96.
- 36: Saebelfeld M, Minguez L, Griebel J, Gessner MO, Wolinska J, 2016. Humic dissolved organic carbon drives oxidative stress and severe fitness impairments in *Daphnia*. *Aquat. Toxicol.* 182: 31-38.
- 37: Pikkemaat MG, Dijk SO, Schouten J, Rapallini M, van Egmond HJ, 2008. A new microbial screening method for the detection of antimicrobial residues in slaughter animals: The Nouws antibiotic test (NAT-screening). *Food Control* 19:781–789.
- 38: Gutierrez MF, Gagneten AM, Paggi JC, 2013. Acute and behavioral sensitivity of *Mesocyclops longisetus* to atrazine and endosulfan formulations under predation pressure. *Water Air Soil Pollut.* 224: 1375-1383.
- 39: LaLone CA, Villeneuve DL, Olmstead AW, Medlock EK, Kahl MD, Jensen KM, Durhan EJ, Makynen EA, Blank CA, 2012. Effects of a glucocorticoid receptor agonist, dexamethasone, on fathead minnow reproduction, growth, and development. *Environ. Toxicol. Chem.* 31: 611-622.

- 40: Escher BI, Baumgartner B, Koller M, Treyer K, Lienent J, McAndell CS, 2011. Environmental toxicology and risk assessment of pharmaceuticals from hospital wastewaters. *Water Research* 45(1):75-92.
- 41: Wu WZ, Li W, Xu Y, Wang JW, 2001. Long-term toxic impact of 2,3,7,8-tetrachlorodibenzo-p-dioxin on the reproduction, sexual differentiation, and development of different life stages of *Gobiocypris rarus* and *Daphnia magna*. *Ecotoxicol. Environ. Saf.* 48: 293-300.
- 42: Pieterse B, Felzel E, Winter R, Van der Burg B, and Brouwer A, 2013. PAH-CALUX, an Optimized Bioassay for AhR-Mediated Hazard Identification of Polycyclic Aromatic Hydrocarbons (PAHs) as Individual Compounds and in Complex Mixtures. *Environ. Sci. Technol.* 47: 11651–11659.
- 43: Ha MH and Choi J, 2009. Effects of environmental contaminants on hemoglobin gene expression in *Daphnia magna*: a potential biomarker for freshwater quality monitoring. *Arch. Environ. Contam. Toxicol.* 57: 330-337.
- 44: OSPAR, 2014. Establishment of a list of Predicted No Effect Concentrations (PNECs) for naturally occurring substances in produced water. OSPAR Agreement 2014-05.
- 45: Lee W, Kang CW, Su CK, Okubo K, and Nagahama Y, 2012. Screening estrogenic activity of environmental contaminants and water samples using a transgenic medaka embryo bioassay. *Chemosphere* 88: 945-952.
- 46: Park S and Choi K, 2008. Hazard assessment of commonly used agricultural antibiotics on aquatic ecosystems. *Ecotoxicology* 17: 526–38.
- 47: Halling-Sorensen, B, 2000. Algal toxicity of antibacterial agents used in intensive farming. *Chemosphere* 40, 731–739.
- 48: Lin T, Chen YQ and Chen W, 2013. Impact of toxicological properties of sulfonamides on the growth of zebrafish embryos in the water. *Environ. Toxicol. Pharmacol.* 36: 1068-1076.
- 49: Nunes B, Carvalho F, Guilhermino L, 2005. Acute toxicity of widely used pharmaceuticals in aquatic species: *Gambusia holbrooki*, *Artemia parthenogenetica* and *Tetraselmis chuii*. *Ecotoxicol. Environ. Saf.* 61: 413–419.
- 50: Yang LH, Ying GG, Su HC, Stauber JL, Adams MS, Binet MT (2008). Growth-inhibiting effects of 12 antibacterial agents and their mixtures on the freshwater microalga *Pseudokirchneriella subcapitata*. *Environ. Toxicol. Chem.* 27: 1201-1208.

- 51: European Union, 2012. Commission Regulation (EU) No 252/2012. Laying down methods of sampling and analysis for the official control of levels of dioxins, dioxin-like PCBs and non-dioxin-like PCBs in certain foodstuffs and repealing Regulation (EC) No 1883/2006
- 52: Bovee TFH, Bor G, Heskamp HH, Hoogenboom LAP, Nielen MWF, 2006. Validation and application of a robust yeast estrogen bioassay for the screening of estrogenic activity in animal feed. *Food Additives and Contaminants* 23: 556-568
- 53: Gizzi G, Hoogenboom LAP, Von Holst C, Rose M, and Anklam E, 2005. Determination of dioxins (PCDDs/PCDFs) and PCBs in food and feed using the DR CALUX® bioassay: Results of an international validation study. *Food Additives & Contaminants*, 22: 472 – 481
- 54: Brand W, De Jongh CM, Van der Linden SC, Mennes W, Puijker LM, Van Leeuwen CJ, Van Wezel AP, Schriks M, and Heringa MB, 2013. Trigger values for investigation of hormonal activity in drinking water and its sources using CALUX bioassays. *Environ. Int.* 55:109-118.
- 55: Jarošová B, Bláha L, Giesy JP, and Hilscherová K, 2014. What level of estrogenic activity determined by in vitro assays in municipal waste waters can be considered as safe? *Environment International* 64: 98–109.
- 56: Kunz PY, Kienle C, Carere M, Homazava N, and Kase R, 2015. In vitro bioassays to screen for endocrine active pharmaceuticals in surface and waste waters. *J. Pharmac. Biomed. Anal.* 106: 107–115.
- 57: Leusch FDL, Khan SJ, Gagnon MM, Quayle P, Trinh T, Coleman H, Rawson C, Chapman HF, Blair P, Nice H, and Reitsema T, 2014. Assessment of wastewater and recycled water quality: A comparison of lines of evidence from in vitro, in vivo and chemical analyses. *Water Research* 50: 420-431.
- 58: Hamers T, Leonards PEG, Legler J, Vethaak AD, and Schippers CA, 2010. Toxicity profiling: An integrated effect-based tool for site-specific sediment quality assessment. *Integrated Environmental Assessment and Management* 6:761–773
- 59: Van Ede KI, Van Duursen MBM, and Van den Berg M, 2016. Evaluation of relative effect potencies (REPs) for dioxin-like compounds to derive systemic or human-specific TEFs to improve human risk assessment. *Arch. Toxicol.* 90: 1293-1305.
- 60: Brack W, Ait-Aissa S, Burgess RM, Busch W, Creusot N, Di Paolo C, Escher BI, Hewitt LM, Hilscherova K, Hollender J, Hollert H, Jonker W, Kool J, Lamoree M, Muschket M, Neumann S, Rostkowski P, Ruttkies C, Schollee J, Schymanski EL, Schulze T, Seiler T-B, Tindall AJ, Umbuzeiro GDA, Vrana B and Krauss M, 2016. Effect-directed analysis

supporting monitoring of aquatic environments - An in-depth overview. *Sci. Total Environ.*, 544: 1073-1118.