STATE OF THE SCIENCE REPORT

GLOBAL WATER RESEARCH COALITION

IN VITRO BIOASSAYS: CURRENT STATUS AND FUTURE APPLICATION FOR WATER MANAGEMENT

Lead agent: Global Water Research Coalition

Prepared by: Shane A. Snyder and Frederic D.L. Leusch

AUGUST 2018

TABLE OF CONTENTS

LIST OF FIGURES

1. THE ISSUE WITH MIXTURES AND "UNKNOWN UNKNOWN" CHEMICALS

Life itself is based primarily upon organic chemistry, which has evolved over the earth's history. Indeed, many organic chemicals of natural origin are necessary for human existence, likewise, many natural organic chemicals are toxic. In fact, what often determines the difference between health benefit and health adversity is the dose and homeostatis (equilibrium) among biological agonists and antagonists (substances that initiate or inhibit a biological response, respectively). This leads to the obvious principle that not all chemical mixtures are toxic nor are all chemical mixtures beneficial. What determines benefit or adverse outcome depends on the chemical composition, dose (exposure) and duration.

Synthetic chemistry began around the mid 1800's with the development of the synthetic dye, "mauve" (Welham 1963). By the start of the 20th century, Germany was producing 90% of the worldwide output of synthetic dyes (Jarman and Ballschmiter 2012). The synthesis of organic dyes led to several discoveries of medicinal properties and the subsequent production of the first synthetic pharmaceuticals, including salicylic acid (aspirin), which in turn gave birth to some of the first pharmaceutical companies at the end of the 19th century (Jones 2011). Thus, aspirin has become an essential global medicine with an estimated consumption of more than 120 billion tablets per year (Warner and Mitchell 2002). Traces of aspirin have now been detected in wastewater discharges (Rabiet, Togola et al. 2006); therefore, aspirin could be considered an emerging contaminant despite more than 100 years of commercial production. The predictions are staggering for synthetic chemical growth (Figure 1), with chemical production quickly outgrowing the comparable human population.

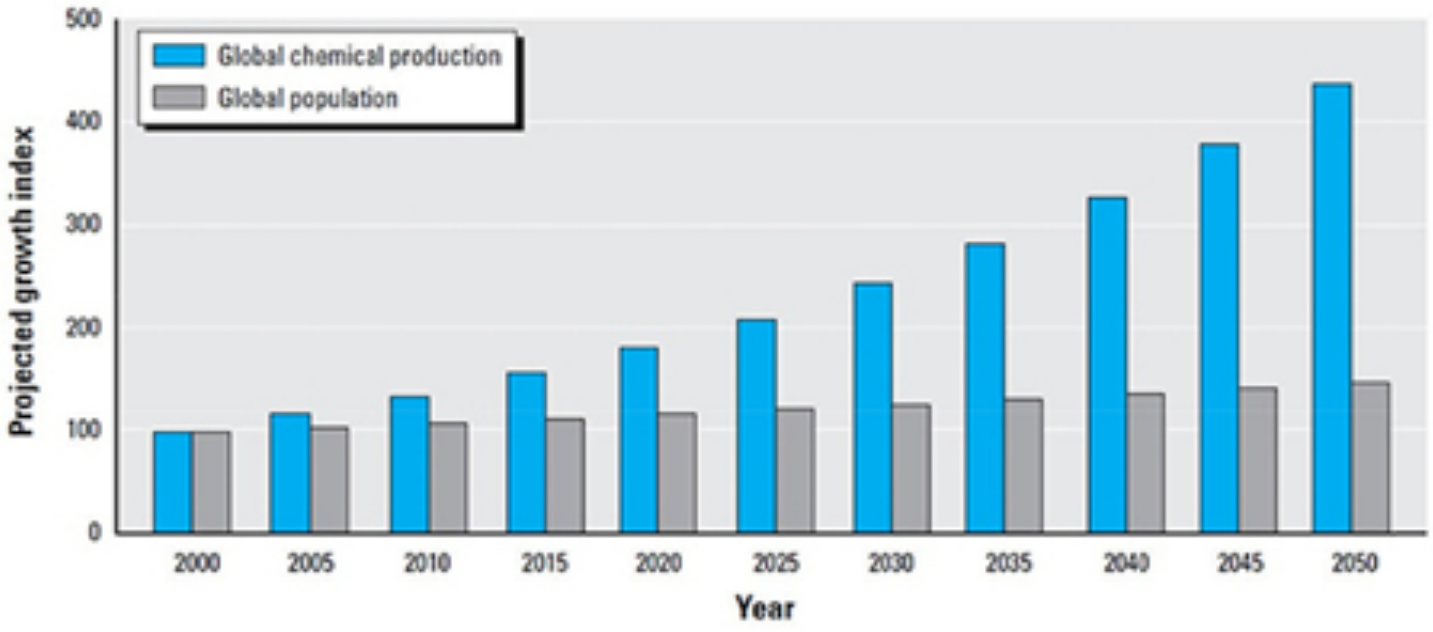

Figure 1. Projected growth in worldwide chemical production and global population through 2050. Source: https://www.ncbi.nlm.nih.gov/books/NBK268889/

There are more than 100,000 chemicals estimated in everyday commercial use (Schwarzenbach et al. 2006; European Chemicals Agency 2017; US Environmental Protection Agency 2017), including 4000 pharmaceuticals alone (Boxall et al. 2012). Once released in the environment, each one of these chemicals can produce a myriad of environmental transformation products via photo-, chemical, and/or bio-transformation. With increasingly sensitive chemical analysis methods (e.g., high resolution mass spectrometry), which now enable detection of chemicals in parts per trillion (ng/L) and lower concentrations, it is becoming increasingly clear that even "pristine" water sources contain a wide range of anthropogenic chemicals from activities upstream or in the catchment. Even rainwater has been shown to contain trace amounts of chemical contaminants (e.g., Hamers et al. 2001), some of which are transported globally through the atmosphere. While these low concentrations are unlikely to pose a significant health concern, there is a scarcity of toxicology information on many of the chemicals currently in commercial use, and in most cases it is impossible to conduct a proper risk assessment for trace chemicals.

In addition, research into mixture toxicity over the past decade has conclusively shown that a mixture of chemicals present at individual concentrations below their respective no-effect (aka guideline) value can still produce a noticeable effect when combined (e.g., Silva et al. 2001). This "something from nothing" is not due to exotic chemical interactive effects such as synergism, which are unlikely to occur at the low chemical concentrations usually found in water samples, but instead due to concentration addition (European Commission 2009). Likewise, some mixtures reduce toxicity of the individual components.

Derivation of chemical guideline values and regulatory standards always includes the use of uncertainty factors (also sometimes referred to as "safety" or "extrapolation" factors) to account for
intra- and inter-species differences, protection of sensitive subpopulations, database uncertainties, and others, which provide an additional level of protection to human consumers (e.g., Ritter et al. 2007). Nevertheless, it is theoretically possible that a water sample that would be compliant with an individual chemical guideline or standard value would nevertheless produce an adverse health effect due to mixture effects.

These two limitations (a near infinite list of analytes and the potential for mixture effects) have led to intense research efforts to apply *in vitro* bioassays, commonly used in the early pre-screening stages of drug development, to water quality assessment (Leusch and Snyder 2015, Escher and Leusch 2012). To highlight the parallel with conventional chemical analysis techniques, these methods are sometimes also referred to as "bioanalytical tools". In cell-based bioanalytical tools, cells of the target organism (humans in the case of drinking water) are used as a surrogate for specific systems

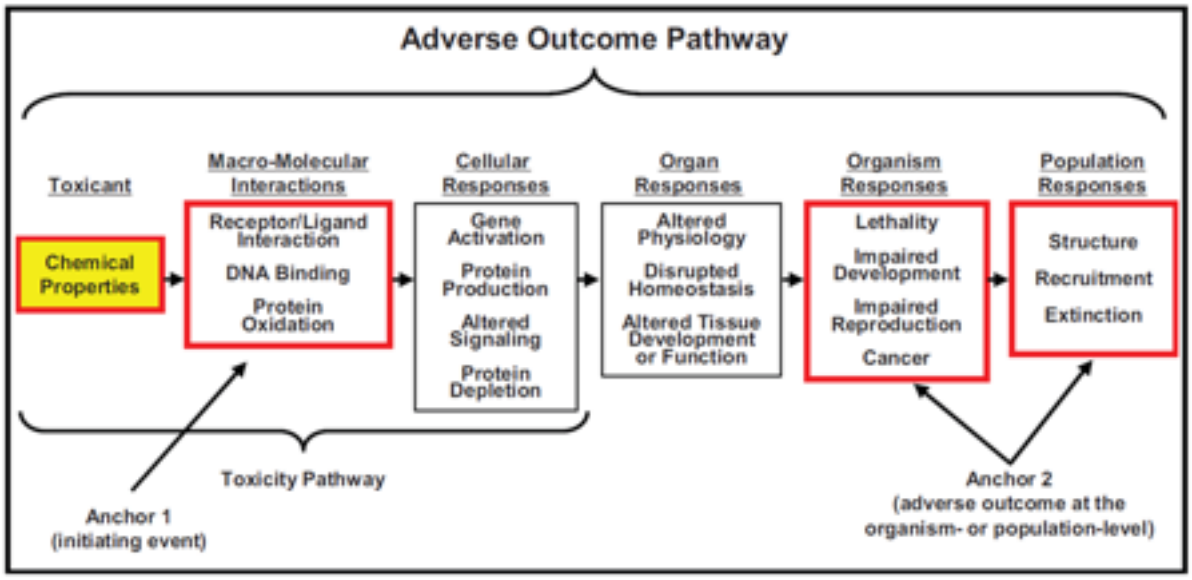

Figure 2. The Adverse Outcome pathway (Ankley et al. 2010).

within an intact organism. In this approach, cells are exposed to substances within a water sample and the cellular response (ranging from a subtle change in gene expression to cell death) is measured after a pre-determined exposure duration (usually 16-24h).

Humans are of course much more complex multicellular organisms and the toxicity to individual cells is not an accurate measure of toxicity at the whole organism level. But advances in molecular toxicology and toxicokinetic modelling in the last decades and well-funded research programs in the US (e.g., Tox21) and Europe (e.g., European Commission Joint Research Centre) have developed a toxicity paradigm termed the Adverse Outcome Pathway (AOP) as the framework to connect effects at the molecular and cellular level ("key initiating event") with effects at the whole organism level ("apical adverse effect"), and have started filling in some of the links ("key events") between those steps (Figure 2).

There is a high level support for this approach by representatives of the National Institutes of Health (NIH), the Food and Drugs Administration (FDA) and the National Research Council (NRC) in the United States (Collins et al 2008, Hamburg 2011, NRC 2007). This approach has also been suggested for evaluation of water disinfection byproducts which have little, if any, in vivo toxicity data (Bull, Reckhow et al. 2011).

The development and adoption of in silico and *in vitro* screening methods is also a recognition of some of the limitations of our current historical toxicity testing paradigm, which relies on whole animal testing. The latter suffers from ethical and financial issues, but is also unable to cope with the sheer number of new chemicals. In addition, historically used animal models also have limitations in the evaluation of potential human toxicity. For instance, there are significant differences between the types and sexes of animals used to predict human toxicity (Bull, Kolisetty et al. 2012). These differences have been discussed in a number of review papers, which highlight the challenge of identifying an appropriate animal model for a variety of human health endpoints (Knight 2007, van Vliet 2011, McGonigle and Ruggeri 2014). A meta-analysis of toxicity data from >2000 drugs tested in four animals (dog, mouse, rat and rabbit) showed that results from animal tests are highly inconsistent predictors of toxic responses in humans (Bailey et al 2014). Thus, even animal testing has limitations and uncertainties in accurately predicting human health relevance.

2. CURRENT APPLICATIONS AND CHALLENGES

2.1. Current applications

Over the past two decades, an increasing number of scientific publications have applied *in vitro* bioassays to water quality assessment, producing a large body of literature and a wealth of information to determine their strength, weaknesses and suitable applications, along with significant improvements in both available assays and protocols. Many *in vitro* bioassays are now mature analytical methods (e.g., Maruya et al. 2016, Schroeder et al. 2016, Wernersson et al. 2015, Altenburger et al. 2015)

In vitro bioassay techniques have been used extensively in assessing recycled water quality (Leusch and Snyder 2015). These studies have shown the suitability of *in vitro* bioassays to assess water quality and demonstrated the added benefit of more comprehensive water quality evaluation. For example, *in vitro* bioassays have been used: (1) to benchmark different water sources, allowing determination of the most suitable source for a specific use (e.g., Rosenmai et al 2018, Escher et al 2014); (2) to assess and monitor water and wastewater treatment efficacy (e.g., Nivala et al 2018, Leusch et al 2014); (3) to detect toxic transformation products produced by water treatment (e.g., Hebert et al 2018, Jia et al 2015); and (4) as a surrogate measure for a wide range of chemical compounds (e.g., Neale et al 2017a, Konig et al 2017, Tang et al 2014). Several studies have also demonstrated the reliability and robustness of *in vitro* bioassay methods in this context (e.g., Di Paolo et al 2016, Jia et al 2015, Mehinto et al 2015, Leusch et al 2010).

2.2. Current limitations

Applying *in vitro* bioassays to water quality testing has some clear advantages, such as the ability to detect chemical contaminants by their effect, the ability to integrate concentration and potency, the ability to quantify some level of mixture effects, and the ability to complement current analytical methods and identify biologically active contaminants in effects-directed analysis (EDA). These advantages make them suitable for a range of purposes, outlined in section 2.1. But there are of course also some limitations of this approach, including issues of sample preparation and quality assurance (see section 2.2.1), how to interpret *in vitro* bioassay results when they do not necessarily imply an adverse effect in exposed whole organisms (see section 2.2.2) and regulatory and end-user concern about how they are applied (see section 3).

2.2.1. Challenges for extraction and identification

The vast majority of *in vitro* bioassay screening of environmental samples involves extraction and concentration of constituents prior to conversion to a compatible solvent and adding a limited amount to the cell culture. The vast majority of applications rely on solid-phase extraction (SPE) followed by solvent elution, evaporation and exchange. This intrinsically leads to limitations in the application of *in vitro* bioassays for comprehensive evaluation of bioactivity. First, most metals and other inorganic salts are not likely to be extracted by typical SPE protocols. For instance, the emerging contaminant perchlorate, a known thyroid endocrine disruptor, would not be captured by most extraction protocols currently applied. Also, highly volatile species are not easily trapped and concentrated, such as trihalomethanes and other volatile organic compounds, although they can contribute to some of the effects (e.g., Stalter et al 2016). Often, specialized methods are required to trap chemicals based on physical-chemical properties and not one extraction/concentration procedure will isolate all chemicals within a complex aqueous mixture. Consequently, multi-layer SPE cartridges with a number of different sorbents to capture chemicals with a wide range of physiochemical properties have recently been applied to extract water samples for bioanalysis (Neale et al. 2017b). However, this limitation applies equally to targeted and non-targeted analytical instrumental approaches. For example, a recent study has shown similar recovery of *in vitro* effect in a range of bioassays by large volume SPE compared to individual recovery for 579 organic micropollutants spiked surface water (Neale et al. 2018).

One important consideration that applies universally to methods used to assess water quality is the requirement for rigorous quality assurance and quality control (QA/QC). Specifically, on the topic of *in vitro* bioassays screening, it

is strongly recommended that at least one positive control substance be spiked into the matrices of interest for each bioassay endpoint evaluated. In the case of *in vitro* estrogenicity, an estrogen agonist should be spiked into the water matrix to be evaluated and the resulting recovery calculated based on a non-spiked duplicate. In addition, method detection and reporting limits for bioasssays should be based on the detection limit of the assay itself as well as the entire method from sample collection to data processing (Winslow, Pepich et al. 2006). Limited information is available on holding times, container types, preservative effects, and other aspects of QA/QC which is relatively more defined for targeted analytical methods (Vanderford, Mawhinney et al. 2011, Vanderford, Drewes et al. 2014). Data on these types of QA/QC as applied to *in vitro* bioassays are relatively sparse and should be addressed before wider application of *in vitro* bioassays in regulatory contexts. Additionally, these types of QA/QC procedures also have not been widely applied to non-targeted instrumental analyses.

Non-targeted analyses (NTA) has also been suggested as a possible option to identify yet unknown chemical constituents in complex aqueous mixtures. However, NTA suffers from the same challenges in extraction/concentration as do *in vitro* bioassay approaches. In addition, NTA generally requires far more expensive instrumentation with an even greater level of expertise needed to operate the instruments and interpret the resulting data. Iterative approaches whereby bioassay data is used in concert with chemical fractionation to better isolate portions of complex chemical mixtures for further characterization offer tremendous promise. This is often referred to as bioassay directed analyses (or effect-directed analysis, EDA) (e.g., Muschket et al 2018, Hashmi et al 2018). Thus, when positive bioactivity is detected, the sample can be fractionated by polarity or molecular weight then each fraction reevaluated using the particular bioassay. In this way, a much more narrow field of potential substances can be isolated and better evaluated for possible identification followed by biological activity – chemical composition balance (Snyder, Villeneuve et al. 2001, Jia, Wu et al. 2016).

2.2.2. Demonstrating the link between *in vitro* effect and in vivo response

There is significant scientific interest in developing a framework to enable the use of *in vitro* bioassays to assess not just the quality, but the safety of drinking water. For example, the BRAVE initiative (http://www.bravebioassays.info) is supported by various water research providers (Water Environment and Reuse Foundation, Global Water Research Coalition, Australian Water Recycling Centre of Excellence, Water Research Australia, PUB Singapore) to develop a research map towards this goal. The think-tank has identified four main areas to focus efforts on, including 1) identifying adverse effect endpoints relevant to water consumption, 2) translating *in vitro* responses to in vivo effects, 3) converting *in vitro* concentration to in vivo exposure dose, and 4) adapting mixture modelling to support each of these steps. While some progress has already been made in some of these areas (e.g., Escher et al 2014, Sonneveld et al 2005, Sonneveld et al 2011, Combes 2012, Benigni 2012, Wetmore 2015, Punt et al 2013), *in vitro* bioassays alone are not capable of assessing the safety of water. But that should not negate the fact that *in vitro* bioassays can expand our analytical universe a little further into the unknown, thereby providing a more comprehensive evaluation of water quality and more rapidly identifying substances which may pose a risk to public health.

3. OUTLOOK FOR BIOASSAY APPLICATIONS

3.1. Overcoming issues with loss of chemicals by sample preparation

While SPE pre-concentration is commonly used when applying *in vitro* bioassays to environmental water samples, the availability of powdered cell culture media means that it is possible to dose un-extracted (filter sterilised) water samples for testing at a relative enrichment factor (REF) close to 1 (an REF of 1 means an undiluted sample, with an REF > 1 indicating a concentrated sample and an REF < 1 a diluted sample). In a recent study, Niss et al. (2018) applied this approach and were able to detect a range of responses in the water samples, including estrogenicity and aryl hydrocarbon activity.

3.2. Development of Effect Based Trigger Values (EBT)

An important aspect in the application of bioassays in water quality assessment is the development of response mechanisms for positive results. In other words, at what level of response should additional actions be taken. One proposed method is the development of benchmark values, which have been termed Effects Based Trigger values (EBT). Various approaches have been proposed to derive EBTs; however, they often produce comparable EBTs for the same endpoints. For example, Brand et al (2013) proposed an EBT of 3.8 ng/L estradiol equivalents (EEQ) for estrogenic activity in drinking water using an Acceptable Daily Intake value for the reference compound estradiol, and applying toxicokinetic parameters to extrapolate *in vitro* activity to in vivo exposure dose. Using a completely different approach, Escher et al (2015) proposed an EBT of 0.2-1.8 ng/L EEQ (depending on the specific bioassay) for estrogenic activity in drinking water using read-across from current chemical guideline values and their relative *in vitro* potencies. Considering that these approaches are based on completely different assumptions (and the usual uncertainties associated with deriving even conventional guideline values), these values are remarkably close. A recent paper by Escher et al. (2018) discusses several options to derive EBTs and provides a large number of proposed EBTs for different *in vitro* and in vivo bioassays to support the European Water Framework Directive.

3.3. Regulatory applications

There are a variety of ways in which *in vitro* bioassays could be used to support drinking water guidelines, ranging from:

- an additional measure of water quality during the initial assessment of suitability of a new water source,
- as a measure of treatment efficacy, either during validation or verification of a new water supply scheme,
- as a routine water quality monitoring tool to identify changes in water characteristics that may trigger further investigation,
- or as a water quality measure with a set numerical target (i.e., guideline or standard)

There is common support for application of *in vitro* bioassays for any of the purposes listed above except the last one (Leusch and Snyder 2015) . The development of numerical *in vitro* bioassay based standards/guidelines would not only be counter to the current application of this science (which has focused on empowering water utilities to better understand the suitability of different water sources for different end uses), but also presume that *in vitro* responses can accurately predict in vivo effects, which is not always plausible as discussed in section 2.2.2.

While in vivo bioassays are already applied in surface and wastewater regulations (e.g., US EPA's Whole Effluent Toxicity (WET) testing requirements, and Direct Toxicity Assessment (DTA) in the Australia-New Zealand Guidelines for Fresh and Marine Water Quality), to the authors' knowledge there are no current regulatory applications of *in vitro* bioassays for water quality assessment. However, numerous expert panels have suggested that *in vitro* bioassays be employed for monitoring of water quality (Drewes et al 2018, WHO 2017, EPHC/NHMRC/NRMMC 2008). In fact, a recent expert panel report from the State of California, USA has indicated that in their expert opinion, bioassays including ER and AhR already are developed enough for immediate implementation as monitoring tools for recycled water (Drewes et al 2018). The California expert panel stated that, "The Panel recommends that the Estrogen Receptor alpha (ER-⊠) and

the Aryl hydrocarbon Receptor (AhR) bioassays be used to respectively assess estrogenic and dioxin-like biological activities in recycled water." Moreover, the panel further stated that those two bioassays were recommended because "each have clear adverse outcome pathways that allow specific molecular responses to be adequately standardized for screening recycled water quality at potable reuse projects." However, the panel recognized that a process to respond to positive results was not yet "sufficiently mature" and thus recommended that implementation be for monitoring/data collection only at this point in time.

Beyond the California expert panel, other step-wise frameworks have been suggested to help water utilities understand how to apply and use the results in bioanalytical tools in water quality monitoring (e.g., Figure 3). All reasonable suggestions so far recommend the use of *in vitro* bioassays to detect unexpected contaminants in addition to available and emerging chemical methods, and not as stand-alone parameters with hard numerical determinants of quality. In addition, the application of *in vitro* bioassays also provides a more comprehensive view of the chemical mixture biological activity, which is not accounted for analytical monitoring. The proposed effects-based trigger values provide a way to anchor bioassay responses to a chemical context, and thus determine when bioassay results warrant further investigation by chemical, technical or operational means.

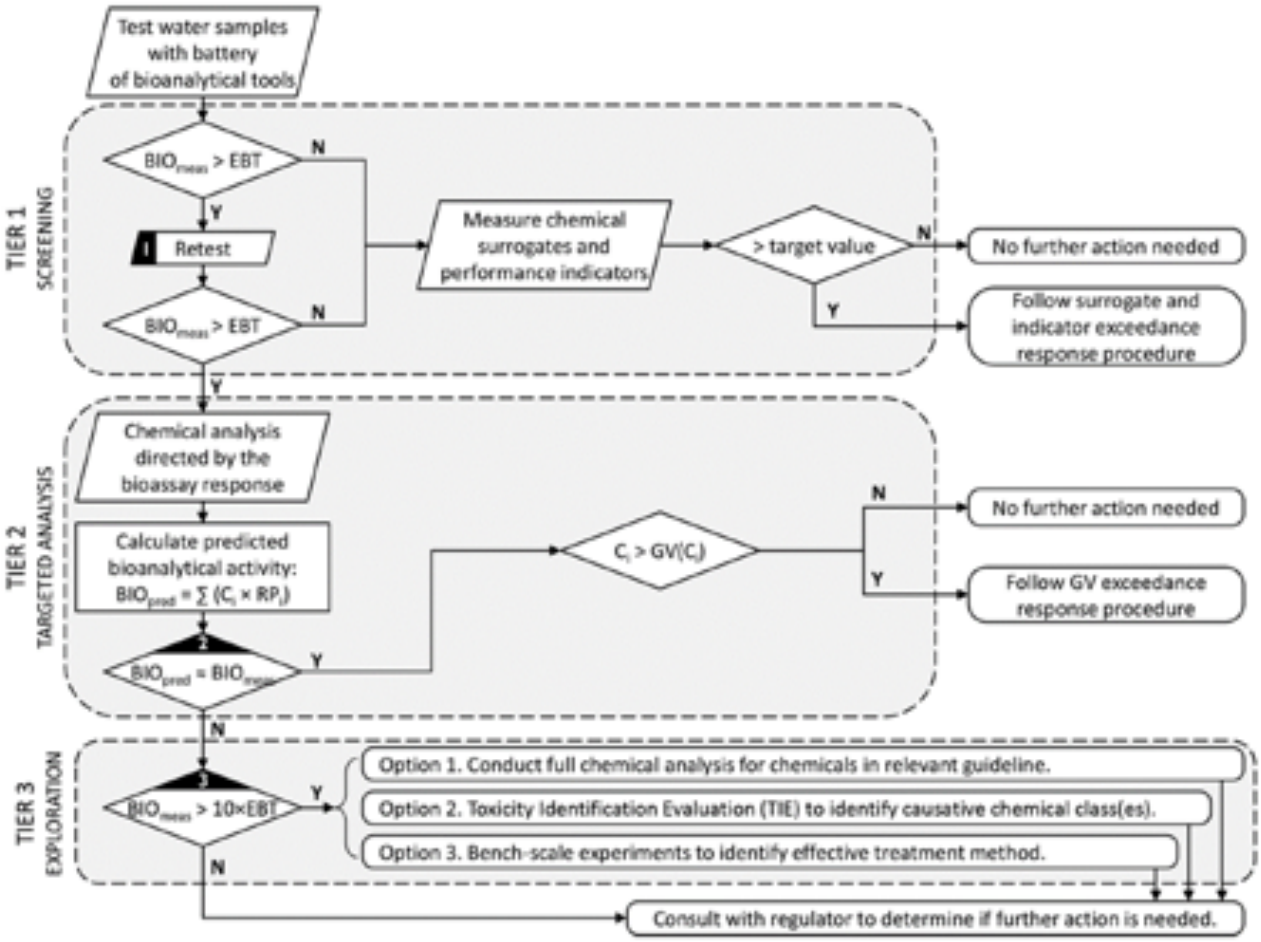

Figure 3. Framework to apply bioanalytical tools in water quality monitoring (Leusch and Snyder 2015).

3.4. Perceptions limiting wider acceptance, and potential solutions

Water agencies and stakeholders have voiced concerns regarding the implementation of *in vitro* bioassays for water quality assessment, including that:

- The use of bioassays will increase the cost of water testing
    - o Compared to most instrumental analyses for trace organic compounds, bioassays require far less costly hardware and methodologies needed are generally familiar to utilities/agencies that conduct microbiological testing.
- The lack of regulatory flexibility to adopt tools when these are not simply "pass" or "fail" measures, but rather "triggers for further investigation"
    - o This will be dependent on the regulatory context. That said, *in vitro* bioassays are already recognized in some advisory and regulatory documents (WHO 2017, EPHC/NHMRC/NRMMC 2008) as useful tools to enhance assessment of water quality.
- "Trigger values" may eventually become adopted as a regulatory compliance measure
    - o There is always a risk that things will get "lost in translation", and it will be critical to ensure that regulators and stakeholders understand the limitations of today's science and do not overreach when applying effects-based trigger values. Further development and acceptance of this field will require ongoing engagement with all stakeholders, including regulators and end-users.
- The methodology and assumptions behind the calculation of effects-based trigger (EBT) values are too difficult to understand
    - o The different methods to calculate EBT currently proposed are all complex, and often mathematically challenging. This is a rapidly evolving field of research, and future developments in this field should consider transparency and ease of use, as well as stakeholder engagement from the start.
- The safety factors included in the calculation of current chemical guidelines are sufficient to address mixture toxicity
    - o The implementation of bioassays for drinking water monitoring should be viewed as an analytical tool to help in the identification of bioactive substances and mixtures that warrant further evaluation. Bioassays can guide analytical methodologies for the identification of specific chemicals in water.
- Testing with *in vitro* bioassays may introduce uncertainty about the safety of drinking water that has already been proven compliant with current regulation
    - o Bioassays are not suggested to replace current guidelines and regulations, but rather, an important tool for providing additional information regarding water quality and more comprehensive analyses to include yet unknown substances.
      While there is always a degree of uncertainty when applying novel technologies, studies that have applied in vitro bioassays to drinking and recycled water have not identified any significant response under normal operations. Conversely, non-targeted analysis (NTA) may identify new substances which have no toxicological data and increase public concern as chemicals detected cannot be described in a public health context.
- Bioanalytical tools will detect activity that cannot be related to a specific chemical, thereby making it difficult to interpret bioassay results
    - o Different types of bioassays will react to different chemicals. Bioassays to measure specific effects (such as receptor-mediated effects) are often responding to a limited scope of compounds (e.g., Tang et al 2014). For example, the natural hormones estradiol, estrone and estriol, the synthetic hormone ethinylestradiol, the natural phytoestrogens genistein and daidzein, and the industrial xeno-estrogens bisphenol A and nonylphenol often account for most if not all of the estrogenic activity in water samples. Other assays that measure non-specific effects (e.g., DNA damage, protein damage, oxidative stress, narcosis) are responding to much larger numbers of chemicals, and often traditionally measured compounds only account for <1% of the overall response. While this indeed poses a challenge for interpretation, this ability to detect a wider range of contaminants is also one of the main drivers for inclusion of *in vitro* bioassays. While the response of these assays cannot be simply translated into a chemical equivalent, there are several proposals to develop meaningful and implementable trigger values (Escher et al 2018) and practical operational responses to bioassay results (Leusch and Snyder 2015).
- Results from *in vitro* bioassays are too variable, and vary between laboratories
    - o Some work has already focused on standardization, including work by the GWRC (GWRC 2008, Leusch et al 2010) and others (Mehinto et al 2015, Di Paolo et al 2015, Hecker et al 2011, Korner et al 2004). These studies have shown that standardization is possible, to a certain extent, and that bioassays are no more variable than chemical analysis of trace contaminants.

4. CONCLUSIONS

In vitro bioassays have been widely used for water quality screening over the past decades, and a large body of literature is now available to evaluate the advantages and limitations of incorporating *in vitro* "effects based assessment" in conventional water quality monitoring. Some of the most mature bioassays are ready to be employed for screening water samples for specific types of toxicity (e.g., estrogens, AhR agonists). The use of bioassays can help build public support by providing more comprehensive screening of unknown water constituents with endpoints based on human health relevance. *In vitro* bioassays cannot determine specific compounds responsible for observed bioactivity and are meant to augment, not replace, existing targeted instrumental assessments. Likewise, non-targeted instrumental analyses without a bioassay component will likely identify substances with no known biological activity, raising additional questions regarding the relevance of those molecules. Taken together, *in vitro* bioassays can guide both targeted and non-targeted instrumental analyses in a process called "effect-directed analysis" (EDA) to identify and prioritise contaminants. While these tools alone are not yet appropriate to determine if water is "safe" or "unsafe", the historical use of animal models to screen the complex mixtures of chemicals in drinking water is vastly infeasible and, as is increasingly recognized, of limited applicability to human health outcomes. Thus, the application of bioanalytical tools for water quality screening is a step forward to increase our understanding of the risks associated with mixtures of chemicals in water.

In a practical context, *in vitro* bioassays could be used for different purposes in water quality assessment, including comparing the quality of different water sources, the efficacy of treatment technologies, and the ultimate chemical quality of drinking water. It is important to keep in mind that a bioassay response does not necessarily suggest an adverse effect in the whole organism, and it is critical for all stakeholders (including regulators, industry and customers) to understand how *in vitro* bioassays can be used. Current industry and regulator concerns focus on the costs of additional bioassay testing, the risk of regulatory creep, the lack of clear interpretation frameworks (despite significant developments in this field over the past 3-5 years), and the unknowns associated with applying a new technology.

It is important to reemphasize that *in vitro* bioassays should be considered as a step forward in more comprehensive water assessment rather than a panacea that will replace analytical monitoring and in vivo animal testing. While limitations surely do exist, *in vitro* bioassays greatly expand our ability to detect contaminants relevant to human and environmental health, and while these tools cannot yet determine whether water is "safe", their application will make water "safer".

5. REFERENCES

Ankley G et al. (2010). Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. Environmental Toxicology and Chemistry 29, 730-741.

Bailey J, Thew M and Balls M (2014) An analysis of the use of animal models in predicting human toxicology and drug safety. Alternatives to Laboratory Animals 42(3), 181-199.

Benigni R (2012) Alternatives to the carcinogenicity bioassay for toxicity prediction: are we there yet? Expert Opinion on Drug Metabolism & Toxicology 8: 407-417.

Boxall ABA et al. (2012) Pharmaceuticals and personal care products in the environment: What are the big questions? Environ. Health Perspect. 120:1221–1229. doi:10.1289/ehp.1104477

Brand W et al. (2013) Trigger values for investigation of hormonal activity in drinking water and its sources using CALUX bioassays. Environment International 55, 109-118.

Bull, R. J., D. A. Reckhow, X. F. Li, A. R. Humpage, C. Joll and S. E. Hrudey (2011). "Potential carcinogenic hazards of non-regulated disinfection by-products: Haloquinones, halo-cyclopentene and cyclohexene derivatives, N-halamines, halonitriles, and heterocyclic amines." Toxicology 286(1-3): 1-19.

Bull, R. J., N. Kolisetty, X. L. Zhang, S. Muralidhara, O. Quinones, K. Y. Lim, Z. X. Guo, J. A. Cotruvo, J. W. Fisher, X. X. Yang, D. Delker, S. A. Snyder and B. S. Cummings (2012). "Absorption and disposition of bromate in F344 rats." Toxicology 300(1-2): 83-91.

Collins FS, Gray GM and Bucher JR (2008) Transforming Environmental Health Protection. Science 319: 906-907.

Combes (2012) Cell Transformation Assays: Are We Barking Up the Wrong Tree? ATLA 40: 115-130.

Di Paolo C et al. (2016) Bioassay battery interlaboratory investigation of emerging contaminants in spiked water extracts–Towards the implementation of bioanalytical monitoring tools in water quality assessment and monitoring. Water Research 104: 473-484.

Drewes JE, Anderson P, Denslow N, Jakubowski W, Olivieri A, Schlenk D and Snyder S (2018) Monitoring Strategies for Constitients of Emerging Concern (CECs) in Recycled Water: Recommendations of a Science Advisory Panel. State Water Resources Control Board, California, USA.

EPHC/NHMRC/NRMMC (2008) Australian Guidelines for Water Recycling: Augmentation of Drinking Water Supplies. Environment Protection and Heritage Council, the National Health and Medical Research Council and the Natural Resource Management Ministerial Council. Canberra, Australia. https://www.nhmrc.gov.au/guidelines-publications/eh56

Escher BI and Leusch FDL (2012) Bioanalytical Tools in Water Quality Assessment. IWA Publishing, London. 266pp.

Escher BI et al. (2014) Benchmarking organic micropollutants in wastewater, recycled water and drinking water with *in vitro* bioassays. Environmental Science and Technology 48, 1940-1956.

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.

Escher et al. (2018) Effect-based trigger values for *in vitro* and in vivo bioassays performed on surface water extracts supporting the environmental quality standards (EQS) of the European Water Framework Directive. Science of the Total Environment 628-629: 748-765.

European Chemicals Agency, 2017. EC Inventory. Retrieved 6 July 2017 from. https://echa.europa.eu/information-on-chemicals/ec-inventory.

European Commission (2009) State of the Art Report on Mixture Toxicity. Report 070307/2007/485103/ETU/D.1.

GWRC (2008) Tools to detect estrogenic activity in environmental waters. Global Water Research Coalition, London, UK. 86pp.

Hamburg MA (2011) Advancing regulatory science. Science 331: 987.

Hamers T, Smit MG, Murk AJ and Koeman JH (2001). Biological and chemical analysis of the toxic potency of pesticides in rainwater. Chemosphere, 45(4-5): 609-624.

Hashmi MAK, Escher BI, Krauss M, Teodorovic I, Brack W (2018) Effect-directed analysis (EDA) of Danube River water sample

receiving untreated municipal wastewater from Novi Sad, Serbia. Science of the Total Environment 624: 1072–1081

Hebert A, Feliers C, Lecarpentier C, Neale PA, Schlichting R, Thibert S and Escher BI (2018) Bioanalytical assessment of adaptive stress responses in drinking water: A predictive tool to differentiate between micropollutants and disinfection by-products. Water Research 132: 340-349

Hecker M et al. (2011). The OECD validation program of the H295R steroidogenesis assay: Phase 3. Final inter-laboratory validation study. Environmental Science and Pollution Research International 18(3): 503-515.

Jarman, W. M. and K. Ballschmiter (2012). "From coal to DDT: the history of the development of the pesticide DDT from synthetic dyes till Silent Spring." Endeavour 36(4): 131-142.

Jia A et al. (2015) *In vitro* bioassays to evaluate complex chemical mixtures in recycled water. Water Research 80: 1-11.

Jia, A., B. I. Escher, F. D. L. Leusch, J. Y. M. Tang, E. Prochazka, B. Dong, E. M. Snyder and S. A. Snyder (2015). "*In vitro* bioassays to evaluate complex chemical mixtures in recycled water." Water Research 80(0): 1-11.

Jia, A., S. Wu, K. D. Daniels and S. A. Snyder (2016). "Balancing the Budget: Accounting for Glucocorticoid Bioactivity and Fate during Water Treatment." Environmental Science & Technology.

Jones, A. W. (2011). "Early drug discovery and the rise of pharmaceutical chemistry." Drug Testing and Analysis 3(6): 337-344.

Knight, A. (2007). "Systematic reviews of animal experiments demonstrate poor human clinical and toxicological utility." Atla-Alternatives to Laboratory Animals 35(6): 641-659.

Konig M et al. (2017) Impact of untreated wastewater on a major European river evaluated with a combination of *in vitro* bioassays and chemical analysis. Environmental Pollution 220: 1220-1230.

Körner W et al. (2004). Interlaboratory Comparison of Four *in Vitro* Assays for Assessing Androgenic and Antiandrogenic Activity of Environmental Chemicals. Environmental Health Perspectives 112: 695-702.

Leusch FDL and Snyder SA (2015) Bioanalytical tools: half a century of application for potable reuse. Environmental Science: Water Research and Technology 1: 606-621.

Leusch FDL et al. (2010) Comparison of five *in vitro* bioassays to measure estrogenic activity in environmental waters. Environmental Science and Technology 44: 3853-3860.

Leusch FDL et al. (2014) Assessment of the application of bioanalytical tools as surrogate measure of chemical contaminants in recycled water. Water Research 49: 300-315.

Maruya KA, Dodder NG, Mehinto AC, Denslow ND, Schlenk D, Snyder SA and Weisberg SB (2016). A tiered, integrated biological and chemical monitoring framework for contaminants of emerging concern in aquatic ecosystems. Integrated Environmental Assessment and Management 12(3): 540-547

McGonigle, P. and B. Ruggeri (2014). "Animal models of human disease: Challenges in enabling translation." Biochemical Pharmacology 87(1): 162-171.

Mehinto AC et al. (2015) Interlaboratory comparison of *in vitro* bioassays for screening of endocrine active chemicals in recycled water. Water Research 83: 303-309.

Muschket M, di Paolo C, Tindall AJ, Touak G, Phan A, Krauss M, et al. (2018) Identification of Unknown Antiandrogenic Compounds in Surface Waters by Effect-Directed Analysis (EDA) Using a Parallel Fractionation Approach. Environmental Science & Technology 52, 288–297.

National Research Council (2007) Toxicity Testing in the 21st Century: A Vision and a Strategy. Committee on Toxicity Testing and Assessment of Environmental Agents, National Research Council. National Academies Press, USA. 217pp.

Neale PA et al. (2017a) Development of a bioanalytical test battery for water quality monitoring: Fingerprinting identified micropollutants and their contribution to effects in surface water. Water Research 123: 734-750.

Neale PA et al. (2017b) Integrating chemical analysis and bioanalysis to evaluate the contribution of wastewater effluent on the micropollutant burden in small streams. Science of the Total Environment 576: 785-795.

Neale PA et al. (2018) Solid-phase extraction as sample preparation of water samples for cell-based and other *in vitro* bioassays. Environmental Science: Processes & Impacts 20(3): 493-504.

Niss F et al. (2018) Toxicity bioassays with concentratred cell culture media – a methodology to overcome the chemical loss by conventional preparation of water samples. Environ Sci Poll Res 25(12): 12183-12188.

Nivala J, Neale PA, Haasis T, Kahl S, König M, Müller RA, Reemtsma T, Schlichting R and Escher BI (2018) Application of cell-based bioassays to evaluate treatment efficacy of conventional and intensified treatment wetlands. Environmental Science: Water Research & Technology 4:206-217

Punt A et al. (2013) Effect of combining *in vitro* estrogenicity data with kinetic characteristics of estrogenic compounds on the in vivo predictive value. Toxicology *in Vitro* 27: 44-51.

Rabiet, M., A. Togola, F. Brissaud, J. L. Seidel, H. Budzinski and F. Elbaz-Poulichet (2006). "Consequences of treated water recycling as regards pharmaceuticals and drugs in surface and ground waters of a medium-sized Mediterranean catchment." Environmental Science & Technology 40(17): 5282-5288.

Ritter L, Totman C, Krishnan K, Carrier R, Vezina A and Morisset V (2007). Deriving uncertainty factors for threshold chemical contaminants in drinking water. Journal of Toxicology and Environmental Health Part B: Critical Reviews 10(7): 527-557.

Rosenmai AK et al. (2018) *In vitro* bioanalysis of drinking water from source to tap. Water Research 139: 272-280.

Schroeder AL, Ankley GT, Houck KA and Villeneuve DL (2016) Environmental surveillance and monitoring—The next frontiers for high‐throughput toxicology. Environmental Toxicology and Chemistry 35(3): 513-525.

Schwarzenbach RP et al. (2006) The challenge of micropollutants in aquatic systems. Science 313:1072–1077. doi:10.1126/science.1127291

Silva E, Rajapakse N and Kortenkamp A (2002) Something from "nothing"– eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environmental Science & Technology, 36(8): 1751-1756

Snyder, S., D. Villeneuve, E. Snyder and J. Giesy (2001). "Identification and quantification of estrogen receptor agonists in wastewater effluents." Environmental Science & Technology 35(18): 3620-3625.

Sonneveld E et al. (2005) Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. Toxicological Sciences 83: 136-148.

Sonneveld E et al. (2011) Validation of *in vitro* screening models for progestagenic activities: Inter-assay comparison and correlation with in vivo activity in rabbits. Toxicology *in Vitro* 25: 545-554.

Stalter et al (2016) Sample Enrichment for Bioanalytical Assessment of Disinfected Drinking Water: Concentrating the Polar, the Volatiles, and the Unknowns. Environmental Science & Technology 50:6495-6505

Tang JYM et al. (2014) Which chemicals drive biological effects in wastewater and recycled water? Water Research 60: 289-299.

U.S. Environmental Protection Agency (2017) Toxic Substance Control Act (TSCA) Chemical Substance Inventory. Retrieved 6 July 2017 from. https://www.epa.gov/tscainventory.

van Vliet, E. (2011). "Current Standing and Future Prospects for the Technologies Proposed to Transform Toxicity Testing in the 21(st) Century." Altex-Alternatives to Animal Experimentation 28(1): 17-44.

Vanderford, B. J., D. B. Mawhinney, R. A. Trenholm, J. C. Zeigler-Holady and S. A. Snyder (2011). "Assessment of sample preservation techniques for pharmaceuticals, personal care products, and steroids in surface and drinking water." Analytical and Bioanalytical Chemistry 399(6): 2227-2234.

Vanderford, B. J., J. E. Drewes, A. Eaton, Y. B. C. Guo, A. Haghani, C. Hoppe-Jones, M. P. Schluesener, S. A. Snyder, T. Ternes and C. J. Woods (2014). "Results of an Inter laboratory Comparison of Analytical Methods for Contaminants of Emerging Concern in Water." Analytical Chemistry 86(1): 774-782.

Warner, T. D. and J. A. Mitchell (2002). "Cyclooxygenase-3 (COX-3): Filling in the gaps toward a COX continuum?" Proceedings of the National Academy of Sciences of the United States of America 99(21): 13371-13373.

Welham, R. D. (1963). "THE EARLY HISTORY OF THE SYNTHETIC DYE INDUSTRY .1. THE CHEMICAL HISTORY." Journal of the Society of Dyers and Colourists 79(3): 98-105.

Wernersson AS et al. (2015). The European technical report on aquatic effect-based monitoring tools under the water framework directive. Environmental Sciences Europe 27(1): 7.

Wetmore B (2015) Quantitative *in vitro*-to-in vivo extrapolation in a high-throughput environment. Toxicology 332: 94-101.

WHO (2017) Potable reuse: Guidance for producing safe drinking-water. World Health Organisation, Geneva, Switzerland. ISBN 978-92-4-151277-0. http://www.who.int/water_sanitation_health/publications/potable-reuse-guidelines/en/

Winslow, S. D., B. V. Pepich, J. J. Martin, G. R. Hallberg, D. J. Munch, C. P. Frebis, E. J. Hedrick and R. A. Krop (2006). "Statistical procedures for determination and verification of minimum reporting levels for drinking water methods." Environmental Science & Technology 40(1): 281-288.Altenburger R et al. (2015) Future water quality monitoring—adapting tools to deal with mixtures of pollutants in water resource management. Science of the Total Environment 512: 540-551.