THE ADVANTAGES OF LINEAR CONCENTRATION-RESPONSE CURVES FOR *IN VITRO* BIOASSAYS WITH ENVIRONMENTAL SAMPLES

Beate I Escher[1,2*], Peta A Neale[3] and Daniel L. Villeneuve[4]

[1]Helmholtz Centre for Environmental Research – UFZ, Permoserstr. 15, 04318 Leipzig;

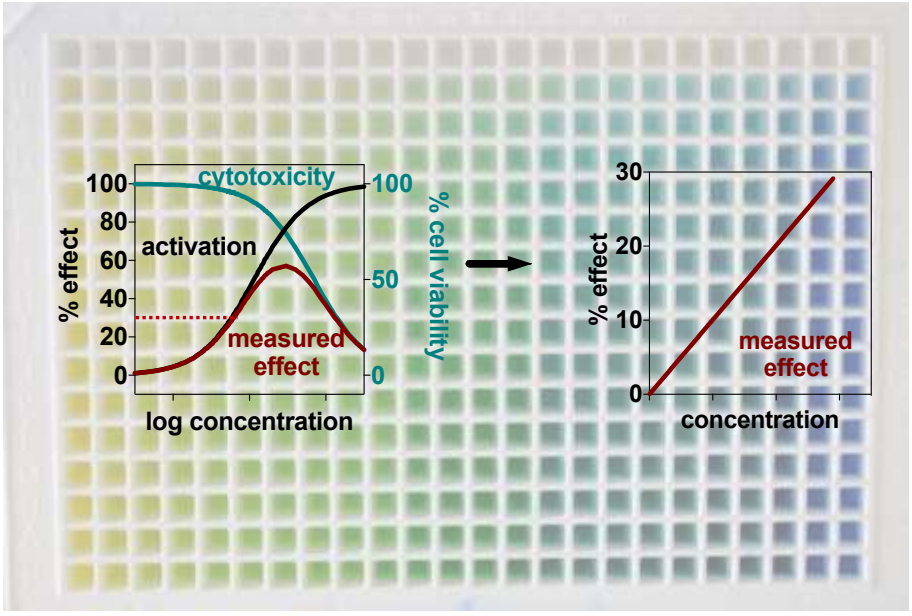
[2]Eberhard Karls University Tübingen, Environmental Toxicology, Center for Applied Geosciences, 72074 Tübingen, Germany

[3]Australian Rivers Institute, School of Environment and Science, Griffith University, Southport QLD 4222, Australia

[4]Mid-Continent Ecology Division, National Health and Environmental Effects Research Laboratory, Office of Research and Development, U.S. Environmental Protection Agency, Duluth, Minnesota, USA

* *Address correspondence to beate.escher@ufz.de; Ph: +49 341 235 1244

Graphical Abstract



Linear concentration-response curves allow simple description of low-dose effects yet rigorous derivation of bioanalytical equivalent concentrations.

Abstract

In vitro assays and high-throughput screening (HTS) tools are increasingly being employed as replacements for animal testing, but most concentration-response curves are still evaluated with models developed for animal testing. We argue that application of *in vitro* assays, particularly reporter gene assays, to environmental samples can benefit from a different approach to concentration-response modelling. First, cytotoxicity often occurs at higher concentrations, especially for weakly acting compounds and in complex environmental mixtures with many components. In these cases, specific effects can be masked by cytotoxicity. Second, for many HTS assays, low effect levels can be precisely quantified due to the low variability of controls in cell-based assays and the opportunity to run many concentrations and replicates when using high density well-plate formats (e.g., 384 or more wells per

plate). Hence we recommend focusing concentration-response modelling on the lower portion of the concentration-response curve, which is approximately linear. Effect concentrations derived from low-effect level linear concentration-response models facilitate simple derivation of relative effect potencies and the correct application of mixture toxicity models in the calculation of bioanalytical equivalent concentrations.

Introduction

Biological data are often log-normally distributed (Limpert 2001) and so are toxicity data. Hence the most popular models for describing sigmoidal concentration-response curves with logarithmic concentration scale (log-CRC) are based on log-normal distributions (probit), log-logistic distributions (logit) and various similar symmetric and asymmetric models that fit deviations from ideal distributions (normal distribution with standard deviation of 1 or a probit CRC with a slope of 1).

One can take different approaches to curve fitting of experimental CRCs for complex environmental samples that contain mixtures of hundreds and thousands of chemicals. The CRC may be described by any best-fit model, even a polynomial, to derive an EC_y value that best describes the concentration triggering the effect y (y being typically 10 or 50% of the maximum effect but also any other measure of effect level of monotonous CRCs) indicated by the absolute residuals in a weighted least squares regression analysis (Scholze 2001). One of the most popular models for continuous data, which has been applied for many decades for microbiological and biochemical assays, is the logistic CRC model using logarithmic concentrations and a four-parameter logistic fit that determines the minimum and the maximum of the

curve as well as the location (EC_{50}) and the slope of the log-linear middle portion of the curve (Volund 1978). Such curves can also be approximated on a linear concentration scale by slope-ratio models at the low and high end of the curve (Finney 1951) or by a log-linear model in the middle portion of the curve (Volund 1978). These models were often adopted in toxicological and ecotoxicological studies that rely on small numbers of observations and replicates to derive precise EC values from logarithmic concentrations (Ritz 2010, Scholze 2001, Slob 2002). The EC values are then typically used for risk assessment where additional extrapolation factors are applied to derive safe concentrations.

For interpretation of high-throughput *in vitro* bioassays data from environmental samples, we rely on data for single chemicals tested with the same bioassay to derive relative (effect) potency (RP or REP) (Villeneuve 2000) to compare effects between chemicals, mixtures and environmental samples. Further we use toxic units (TU = 1/EC) and bioanalytical equivalent concentrations (BEQ) to describe effects of complex samples (Wagner 2013). These approaches require parallel log-CRCs and the same minima and maxima of effects (efficacy) (Villeneuve 2000). The easiest way to comply with the condition for effect-level independent REP is to use only the low effect-level portion of the CRC, where the slope-ratio models apply (Volund 1978). This also tends to be the relevant concentration range for single chemicals in complex environmental mixtures, where the individual components are often present at concentrations below visible effect levels. The CRC on a linear concentration scale (linear-CRC) are typically linear up to 30% of the maximum effect level.

High-throughput screening (HTS) in vitro *bioassays*

Classic whole animal testing generally employs a low number of replicates, often with high variability of the biological controls. Hence, effect levels below 10% of the maximum effect are often not statistically distinguishable from background. This is one of the reasons why no observable effect concentration/lowest observable effect concentration (NOEC/LOEC) and no observable adverse effect level/lowest observable adverse effect level (NOAEL/LOAEL) values were traditionally derived by hypothesis testing (Fox 2016, Green 2013). Benchmark effect concentrations, such as EC_{10}, estimated from dose-response modelling, were advocated when continuous dose-response (or concentration-response) data became more widely available.

The situation is different for many modern *in vitro* assays that can be run in 384- or even 1536-well plate format allowing large numbers of dose levels in one experiment. In addition, precision pipetting with robotic instrumentation allows not only efficient and accurate pipetting with few technical replicates but also the ability to design the spacing of concentrations at will, easily covering anything from several orders of magnitude of serial dilutions to narrow linear spacing. Consequently, repeatability of many cell-based bioassays has become so good that the limit of detection (effect of control wells plus three times standard deviation of controls) is typically below an induction ratio (IR) of 1.5 or an effect level of 10%. This allows the derivation of effect concentrations at low effect levels with high precision and accuracy. Hence, in the case of HTS bioassays, there are not the traditional experimental limitations to working with low level linear-CRCs. Further, unlike for more complex CRC models that will not yield a reliable fit with only partial data or are over-parameterized in the case of partial CRCs, any deviation from linearity is easily visible.

Even more important to consider is the make-up of reporter gene assays, many of which are used in both large chemical screening programs such as Tox21 (Shukla 2010) and in environmental monitoring (Escher 2012a). Reporter gene assays work on the principle that a gene is engineered to be under control of the response element of a specific nuclear receptor or transcription factor and translates into a protein, which can be quantified in a simple way, such as green fluorescence protein or an enzyme that can be quantified by substrate turnover (e.g., luciferase after addition of luciferin and ATP, or β-lactamase after addition of a fluorescent substrate). The reporter protein read out is proportional to the activation of the targeted nuclear receptor or transcription factor. This proportionality will only be valid provided the cells remain viable and are not compromised by cytotoxicity.

How to deal with cytotoxicity in reporter gene assays?

The activation of the reporter gene can be masked by cytotoxicity. This problem is especially pronounced for complex environmental samples where the majority of chemicals in the mixture might have non-specific effects that contribute to cytotoxicity while only a small fraction specifically activates the response of interest. Consequently it is recommended to run parallel cytotoxicity assays and use only non-cytotoxic concentrations when evaluating CRCs.

Even if the cells are not dead, a burst of activity across a wide range of stress response pathways and even nuclear receptors, not directly related to cell death, has been widely observed during screening of single chemicals in the Tox21 program (Judson 2016). Similar effects have been observed when screening environmental samples using cell-based assays (Hebert 2018, Nivala 2018). This phenomenon, termed the cytotoxic burst (Judson 2016), represents another artefact that can

complicate CRC modelling when working with concentrated samples required to elicit higher effect levels.

The current strategy to deal with this phenomenon is to omit any effect data associated with concentrations above the inhibitory concentration leading to 10% reduced cell viability (IC_{10}) (Neale 2017b, Nivala 2018). Using linear CRCs focused on effect levels up to 30% maximum effect only, one greatly reduces the risk of reaching cytotoxic concentrations but it remains important that any concentrations above the IC_{10} for cytotoxicity are omitted in the derivation of EC values for the specific effects. This might lead in some cases to the inability to derive EC values for activation of a specific response. In our experience this situation is quite common for certain assays (e.g., p53) where effects of environmental samples start to occur right around the IC_{10} with often just one or two measured concentrations in the range between EC_y and IC_{10}. In this case we recommend a conservative approach and would report no activity up to cytotoxicity to assure that there are no false-positive results.

Linear concentration-response modelling: why and how?

Cumulative normal distributions of log-CRCs with standard deviations of one, probit log-CRCs with slopes of one or logistic log-CRC with slopes around 1.2 lead to log-concentration-response plots of the sigmoidal form depicted in Figure 1 (dotted lines). Concentrations with any units (e.g., molar, ppb, or mg/L), including relative enrichment factors (REFs) in the case of environmental samples, can be used. A REF indicates the factor by which a water sample would need to be concentrated to achieve the effect level y. For example, a REF of 8 would indicate that 8-fold enrichment would be needed to yield a response equal to 10% of the maximum effect

(y for this example) (Escher 2012a). Such cumulative normal distributions with a standard deviation of one have, per definition, parallel slopes (Figure 1). Reformatting the CRC to a linear concentration scale demonstrates that all CRCs are approximately linear up to an effect level of around 30%. Linear CRCs at lower effect levels are not only less prone to be affected by cytotoxicity but can also be treated mathematically with much simpler models.

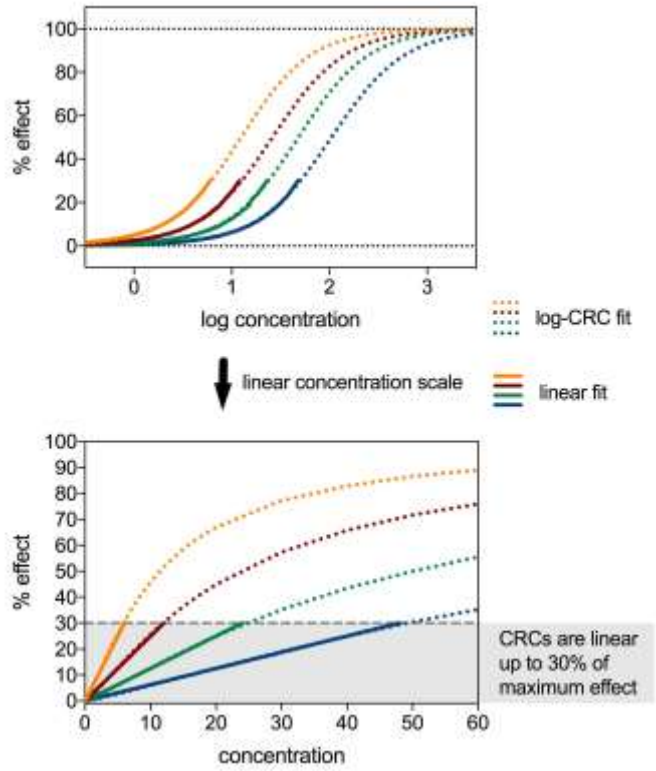


Figure 1. *Relationship between sigmoidal log-CRCs and linear CRCs*

Low-level linear CRCs for HTS in vitro *bioassays*

For reporter gene assays that target the activation of nuclear receptors, the response is typically expressed as relative luminescence units, relative fluorescence units, etc. depending on the mode of detection. These values must be normalized for each microtiter plate by comparison with a known reference compound as a positive

control activating 100% of the defined maximal response and negative controls, i.e., the signal of the unexposed cells or the solvent control, set to 0% response.

The effect concentration triggering effect y (y = 10% or any other benchmark) is defined by eq. 1 and the associated standard error (SE_{ECy}) are derived by error propagation shown in eq. 2. The same equation also holds for cytotoxicity and cell population growth rate and biomass, provided the data can be assigned to effects between zero and 100%.

$$EC_y = \frac{y}{slope} \quad (1)$$

$$SE\ (EC_y) \approx \frac{y}{slope^2} \cdot SE(slope) \quad (2)$$

Super-induction (i.e., responses exceeding the maximum observed for the positive reference compound (Baston 2011)) and partial agonism (i.e., levelling off of the maximum effect at lower than 100% of a full agonist (Howard 2010)) have been observed in some reporter gene assays on nuclear receptors for chemicals and environmental mixtures. Mixture models, like the generalized concentration addition model (Howard 2009), have been developed to account for mixture effects even in complex environmental samples. However, applying these models generally requires that that full CRCs are available for all independent components of the mixture, whether they are full or just partial agonists (Brinkmann 2018). Unfortunately, in complex environmental samples even if some of the composition is known, the complete composition is generally unknown and full CRCs, even for the known components are often unavailable to support the mixture modeling. By applying linear CRCs focused on effect levels below 30%, we can circumvent the problems of partial agonism and super-induction.

There are reporter gene assays, such as those indicative of transcription factors of adaptive stress responses or some genotoxicity assays such as umuC, where there is no maximum (100%) response. In these cases an induction ratio (IR) can be used as a measure of effect. The IR is defined as the ratio of the signal of the sample divided by the signal of the control. The resulting CRCs are typically linear up to an IR of 4 to 5 and IR 1.5 is typically a suitable effect benchmark that is above the limit of detection, which is defined as three times the standard deviation of the effects of the unexposed cells (Buchinger 2010, Escher 2012b, ISO13829:2000 2000). The $EC_{IR1.5}$ is thus derived by eq. 3 and the standard error of the $EC_{IR1.5}$ ($SE(EC_{IR1.5})$) can be calculated by error propagation with eq. 4.

$$EC_{IR1.5} = \frac{0.5}{slope} \quad (3)$$

$$SE\,(EC_{IR1.5}) \approx \frac{0.5}{slope^2} \cdot SE(slope) \quad (4)$$

Examples of applications

To illustrate how these approaches work in practice several examples from water quality assessment are provided (Figure 2).

The first example shows responses for a solid phase extract of a wastewater treatment plant effluent (Figure 2A). Solid phase extraction (SPE) is frequently used as a technique to concentrate micropollutants present in a water sample while separating those compounds from other matrix constituents that can interfere with downstream analyses (Neale 2018). In this example, the activation of the arylhydrocarbon receptor in the AhR CALUX assay (Brennan 2015) by a wastewater plant treatment effluent extract (Nivala 2018) is compromised by cytotoxicity which begins to occur around a REF of 10 (Figure 2A, left). At concentrations that are

already cytotoxic super-induction compared to the maximum response of the 2,3,7,8-TCDD reference compound can be observed. This is clearly an artefact of the cytotoxicity burst (Judson 2016), which does not indicate a specific effect but is a consequence of non-specific toxicity. Fitting a specific effect beyond cytotoxicity would not be mechanistically meaningful and could confuse mixture modelling. These issues are avoided when applying the linear evaluation of the activation-CRC (EC_{10}, Figure 2A, middle) because concentrations causing 10% or more cytotoxicity (IC_{10}, Figure 2A, right) only occurred at concentrations causing more than 30% effect and thus were not included in the analysis.

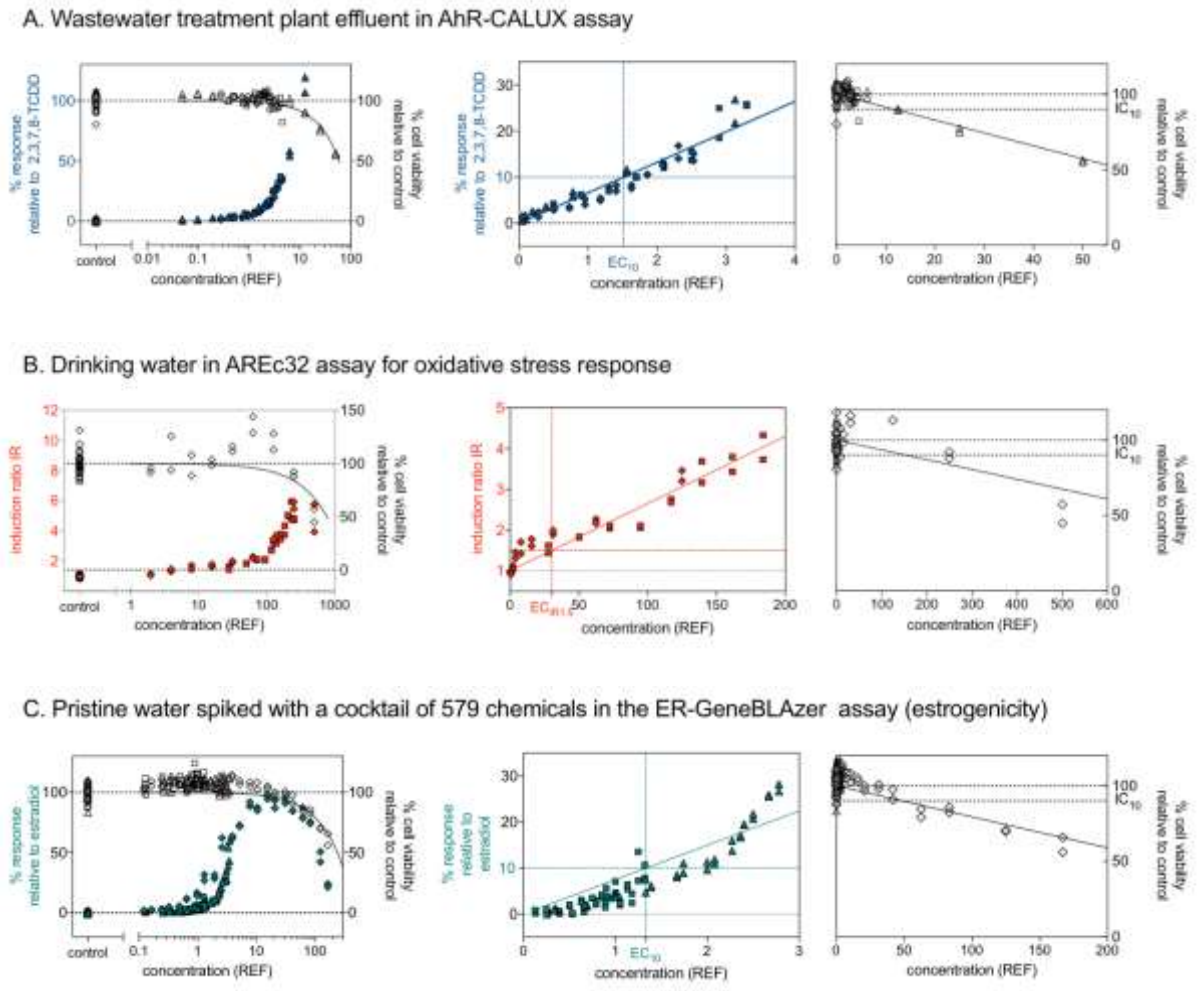


Figure 2. *Examples for the proposed linear CRC evaluation. A. Wastewater treatment plant effluent enriched with SPE, run in the AhR CALUX assay for*

activation of the arylhydrocarbon receptor (data from Nivala 2018). B. Drinking water enriched with SPE, run in the AREc32 assay for oxidative stress response (data from Hebert 2018). C. 579 chemicals spiked to a pristine creek water sample, run in the ER-GeneBLAzer assay for estrogenicity (Neale 2018). The empty symbols are cell viability data and the filled symbols activity data with different symbols from different independent experiments and the same symbols in activity and cytotoxicity from a matching experiment.

A second example refers to activation of the oxidative stress response quantified with the AREc32 assay in drinking water (Figure 2B) (Hebert 2018). Here cytotoxicity was observed at a 200 fold enrichment of the water sample, but a linear CRC allows the derivation of an $EC_{IR1.5}$ of 30 fold enrichment for the oxidative stress response (Figure 2B). Despite the high enrichment needed, the response was not compromised by cytotoxicity at effect levels up to IR 4 and was distinctly different from the control. Hence an $EC_{IR1.5}$ could be derived despite the inverse U-shaped form of the raw CRCs for the activation of oxidative stress response.

The third example refers to a study where relatively clean creek water was spiked with 579 different micropollutants, among them steroidal estrogens but also others that would disturb the estrogenic effect by causing cytotoxicity (Neale 2018). As the non-estrogenic chemicals had a strong effect on cell viability with an IC_{10} of REF 50, the activation of the estrogen receptor quantified with the ER-GeneBLAzer assay (Neale 2018) followed an inverted U-shaped curve but was again fairly linear at low effect levels (Figure 2C). Note that in this example (Figure 2C) a slight deviation from linearity was evident (corresponding to a slope > 1.2 of the log-logistic fit) but the EC_{10} was still within a factor of two from the concentrations corresponding to the experimental IR 1.5, which would equate to 0.3 log-units error as compared to a

perfect sigmoidal fit. This low uncertainty is acceptable because a log-logistic fit would have led to unknown uncertainty for the subsequent analysis of relative effect potency, as discussed in the next chapter.

Relative (Effect) Potency

If we want to compare effects of chemicals between each other and between different bioassay or predict mixture effects of defined chemical mixtures, we need to know the relative potency RP_i of compound i, also called relative effect potency REP_i. REP_i is also a vital parameter for the derivation of effect-based trigger values (Escher 2018, Escher 2015) that are used to define acceptable water quality with respect to mixtures in a similar way as environmental quality standards or guideline values do for single compounds.

REP_i can be calculated by eq. 5 and its associated SE by eq. 6. Using the low-level linear CRCs described above, the REP_i and its SE can also be directly calculated from the slopes (eq. 5) and hence REP_i is independent of the effect level within the linear low-level effect range (Figure 3).

$$REP_i = \frac{EC_y\,(\text{reference})}{EC_y\,(i)} = \frac{\text{slope}\,(i)}{\text{slope}\,(\text{reference})} \quad (5)$$

$$SE(REP_i) \approx \sqrt{\frac{1}{EC_y(i)^2}\cdot SE(EC_y(\text{reference}))^2 + \frac{EC_y(\text{reference})^2}{EC_y(i)^4}\cdot SE(EC_y(i))^2}$$

$$= \sqrt{\frac{1}{\text{slope}(\text{reference})^2}\cdot SE(\text{slope}(i))^2 + \frac{\text{slope}(i)^2}{\text{slope}(\text{reference})^4}\cdot SE(\text{slope}(\text{reference}))^2} \quad (6)$$

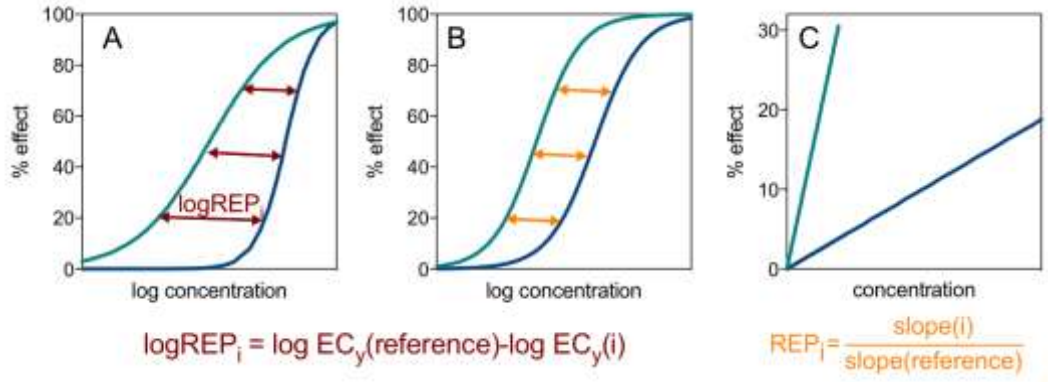


Figure 3. *If the slopes of the log-CRCs are not the same for reference compound and chemical i, then the REP$_i$ are dependent on the effect level (A). For log-CRC with the same slope (B) and linear-CRCs (C), the REP$_i$ are independent of the effect level.*

Another advantage of linear low-level CRC for deriving REP$_i$ is that the concept also works for weakly potent chemicals (partial agonist) that do not reach 50% effect. There are other ways to cope with non-similar concentration effect curves (Dinse 2011) and derive REPs as a function of response level (Ritz 2006) but the linear method described here is probably the most simple for evaluation of HTS data and circumvents complex approaches to derive the REP$_i$ that then cannot be easily applied for the determination of BEQ, which are essential for the description of mixture effects of environmental samples.

Bioanalytical equivalent concentrations

BEQ are commonly used to express the potency of a complex mixture in terms of an equivalent concentration of a well-defined reference compound that produces the same biological response. BEQ_{bio} can be calculated directly from bioassay EC values and BEQ_{chem} are calculated from chemical concentrations and the REPs of all components of a mixture. Comparison of BEQ_{bio} and BEQ_{chem} can be used to determine which fraction of the mixture effect of a water sample is triggered by known and by unknown chemicals (Tang 2014), an approach that is termed “iceberg modelling” (Neale 2018).

BEQ derived from bioassay results (BEQ_{bio})

The calculation of BEQ_{bio} is very similar to the REP_i estimation described above. Because the mixtures are generally undefined, REP_i cannot be calculated directly from a ratio of concentration units, rather an equivalent biological activity is used to associate an enrichment factor, equivalent volume of water, equivalent mass of sediment, etc. with a certain concentration of a reference chemical. There is a fairly confusing nomenclature on BEQ in the literature. For example, scientists often referred to toxic equivalent concentrations (TEQ). However, because many *in vitro* assays are not reporting toxicity but some defined biological effect, bio-equivalents or bioanalytical equivalent concentrations BEQ have become more popular for *in vitro* bioassays. Wagner et al. (2013) reviewed 234 peer-reviewed publications on that topic and came to a very similar conclusion as Villeneuve et al. (2000) did for the REP_i, i.e., that non-parallel log-CRCs and the variability of the maximum effect were the largest impediments, as well as the extrapolation to untested enrichment factors. After this analysis, Wagner et al. (2013) recommended the definition of bio-

equivalents from non-linear interpolations of log-CRC and proposed a checklist to assess the validity of the approach in practical applications when often only single point estimates were available. At the time of that publication the use of linear CRC for environmental samples was already emerging (Escher 2012b, Escher 2013) but they were not accounted for in that review. In recent years the linear approaches and slope-ratio for first developed and applied to single compounds decades ago have been increasingly, and effectively, applied to environmental samples (König 2017, Neale 2015, Neale 2017b, Nivala 2018). However, their advantages and differences relative to other curve fitting and effect concentration estimation approaches have not been thoroughly discussed preventing wider implementation.

As the BEQ_{bio} from effect concentrations in a bioassay is the ratio of the EC_y of a bioassay-specific reference compound divided by the EC_y of the sample (eq. 7), it can be directly calculated from the inverse ratio of the slopes, provided there is an equal intercept of both lines. The associated standard error of the BEQ ($SE(BEQ_{bio})$) can be simply calculated by error propagation (eq. 8). This is a great advantage over more complex models where advanced statistical tools or Monte Carlo resampling methods are required to estimate the uncertainty of the BEQ.

$$BEQ_{bio} = \frac{EC_y\,(reference)}{EC_y\,(sample)} = \frac{slope\,(sample)}{slope\,(reference)} \quad (7)$$

$$SE(BEQ_{bio}) \approx \sqrt{\frac{1}{EC_y(sample)^2} \cdot SE(EC_y(reference))^2 + \frac{EC_y(reference)^2}{EC_y(sample)^4} \cdot SE(EC_y(sample))^2} \quad (8)$$

Naturally the BEQ can also be calculated with the same equation (eq. 7) for EC values that are derived from logistic fits of log-CRCs. However, if the slopes of the logistic log-CRC were not equal for the reference compound and sample, the error

associated with the BEQ would be more difficult to estimate, e.g., by non-linear interpolation yielding a range of BEQ values instead of a single estimate (Schmitt 2012). Approaches to deal with non-linear CRC and quality control measures were discussed in detail by Wagner et al. (2013) giving some guidance for the use of historic data.

In contrast, if the linear-CRC did not yield a perfect linear fit, which would be equivalent to a slope different from 1 of a logistic log-CRC, we can still quantify this deviation by the SE of the slope of the linear CRC and propagate the error all the way through to the BEQ. This is demonstrated by the fact that in the formulation of eq. 7 we do not even need the EC_y values but the BEQ can be directly calculated from the inverse ratio of the slopes and accordingly the SE could also be calculated from the slopes and their errors alone.

BEQ derived from chemical analysis BEQ_{chem}

The BEQ_{chem} can be calculated as the sum of the product of the REP_i and the concentration C_i of all detected chemicals i (eq. 9, SE, eq., 10).

$$BEQ_{chem} = \sum_{i=1}^{n} REP_i \cdot C_i \quad (9)$$

$$SE(BEQ_{chem}) \approx \sqrt{\sum_{i=1}^{n} C_i^2 \cdot SE(REP_i)^2 + REP_i^2 \cdot SE(C_i)^2} \quad (10)$$

The comparison of BEQ_{bio} with BEQ_{chem} can be used as a “mass balance” or “potency balance” analysis if all causative agents are known, e.g. for effect-directed analysis (Hashmi 2018). It can also be used to evaluate whether the known composition of an environmental sample can reasonably account for the biological

activity observed, or whether unknown constituents and/or complex (greater than additive) interactions are likely contributing (Neale 2017a, Tang 2014).

Advantages and limitations of linear CRCs

Overall, the low-dose linear CRC approach proposed here has many practical advantages for application of HTS assays with environmental samples but also one main theoretical caveat. The caveat is that a normal distribution with a standard deviation of one of the effect data or a log-logistic fit with a slope of 1.2 is a prerequisite for the CRC to be linear from 0% to 30% effect level. Based on several years of experience with this linear approach we are confident that the majority of linear-CRCs are fairly linear below effect levels of 30% of maximum effect or IR 4 for a wide range of water, sediment and biota samples but there are exceptions, as shown in Figure 2C. Fortunately, even small deviations from linearity can be easily detected by visual inspection and described by the regression coefficient and the SE of the slope of the linear regression. Generally speaking, complex dose-response modelling is unnecessary and linear fits can easily be accomplished in standard spreadsheet calculation programs. This can both accelerate evaluation of HTS data and help avoid error when non-experts are applying these tools. Additionally, by focusing the analysis on low effect levels, solubility problems are often circumvented, and cytotoxicity interferences can be much better managed, avoiding experimental artefacts.

The major practical limitation is that the serial dilutions which are most commonly used because they can easily be prepared manually, are not perfectly suited for linear CRC modelling. However, the increasingly widespread availability of simple automated dispensers and HTS bioassay robots can help overcome this practical

problem. Furthermore, even using manual dilution approaches, linear dilution series can be prepared.

CONCLUSIONS

In working with users of bioassay data who have different levels of expertise, e.g., regulators, students, scientists from other fields, we have found that a simple and consistent evaluation method is much less prone to error than a complex model where the user must make decisions about the selection of valid data points to be included in the analysis and the models to be applied. Summarizing all discussion points above, we recommend using the entire data set of concentrations against cell viability obtained from testing a given environmental sample to derive the IC_{10} for cell viability (Figure 4A). Then only concentrations below the IC_{10} should be used for further processing (Figure 4B). The remaining data should be visually inspected before plotting concentrations against activity and applying an appropriate concentration-response model to derive the EC_{10} or $EC_{IR1.5}$ for reporter gene activation (Figure 4C). In our experience, in most practical cases we were left with only a low effect level portion of the curve up to 30% (Figure 1), which was often very close to linearity, making linear fits of linear-CRCs (i.e., slope ratio) most amenable for the data analysis. We have successfully applied the low level linear-CRC approach in numerous case studies with over one hundred *in vitro* bioassays applied to many different types of water samples from sewage to surface water to drinking water in collaboration with more than 20 international research groups (e.g., Escher 2014, Neale 2015, Neale 2017a, Neale 2017b, Nivala 2018). However, this approach will attain its full potential only once cytotoxicity is measured in parallel, which was admittedly not the case in all of the previous case studies.

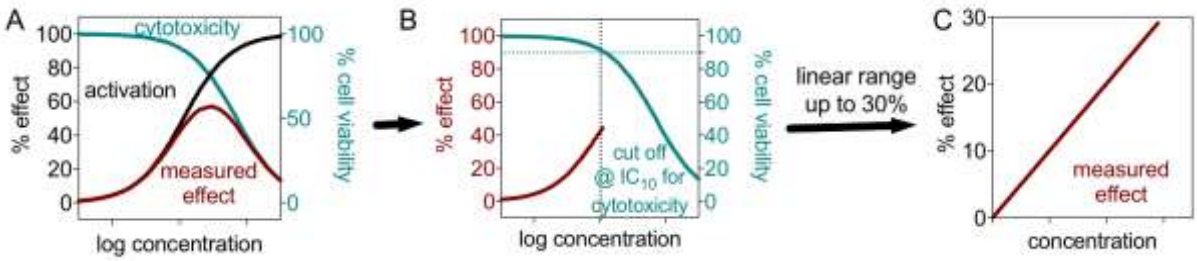


Figure 4. *Recommended processing of CRC data of environmental samples. A. Measured effect yields often U-shaped CRCs due to cytotoxicity overlaying activation. B. All concentrations above the IC_{10} for cytotoxicity should be removed for analysis of effect. C. The linear-CRC model should be only applied to data <30% effect (linear range, see Figure 1).*

The simple and transparent approach with all uncertainty quantified is a good starting point for the regulatory acceptance of *in vitro* bioassays for water quality assessment. Since proposed effect-based trigger (EBT) values for water quality are typically derived as EBT-BEQ (Brand 2013, Escher 2015, van der Oost 2017, Escher 2018) the discussed uncertainties in BEQ from log-CRC modelling (Wagner 2013) would potentially lead to low-quality EBTs and subsequently to high uncertainty in the water quality monitoring and compliance assessment. Furthermore, the simplicity of the linear CRC approach opens up application in sediment and biota testing. Of course, this aspect is not the only one to consider when applying environmental samples in *in vitro* HTS assays (Windal 2005), but it is a crucial one and one can truly simplify life by simplifying modeling.

Acknowledgements

We thank Frederic Leusch, Griffith University, Australia, Etienne Vermeirssen, Swiss Centre for Applied Ecotoxicology, Switzerland, and Sebastian Buchinger, Federal Institute for Hydrology, Germany, for helpful discussions on dose-response modelling and BEQ. We thank Rita Schlichting, Annika Jahnke and Sebastian Buchinger for reviewing the manuscript. We appreciate valuable input from four anonymous reviewers. The presented research was financially supported through the EU-FP7 collaborative project SOLUTIONS (grant agreement no. 603437).

Disclaimer— All authors have no interest to declare. The views expressed in the present review are solely those of the authors.

REFERENCES

Baston DS, Denison MS. 2011. Considerations for potency equivalent calculations in the Ah receptor-based CALUX bioassay: Normalization of superinduction results for improved sample potency estimation. *Talanta* 83:1415-1421.

Brand W, De Jongh CM, van Linden SC, Mennes W, Puijker LM, van Leeuwen C, van Wezel A, Schriks M, Heringa M. 2013. Trigger values for investigation of hormonal activity in drinking water and its sources using CALUX bioassays. *Environ Internat* 55:109-118.

Brennan JC, He G, Tsutsumi T, Zhao J, Wirth E, Fulton MH, Denison MS. 2015. Development of Species-Specific Ah Receptor-Responsive Third Generation CALUX Cell Lines with Enhanced Responsiveness and Improved Detection Limits. *Environ Sci Technol* 49:11903-11912.

Brinkmann M, Hecker M, Giesy JP, Jones PD, Ratte HT, Hollert H, Preuss TG. 2018. Generalized concentration addition accurately predicts estrogenic potentials of mixtures and environmental samples containing partial agonists. *Toxicol in Vitro* 46:294-303.

478 Buchinger S, Grill P, Morosow V, Ben-Yoav H, Shacham-Diamand Y, Biran A,
479 Pedahzur R, Belkin S, Reifferscheid G. 2010. Evaluation of chrono-
480 amperometric signal detection for the analysis of genotoxicity by a whole cell
481 biosensor. *Anal Chim Act* 659:122-128.
482 Dinse GE, Umbach DM. 2011. Characterizing non-constant relative potency. *Regul*
483 *Toxicol Pharmacol* 60:342-353.
484 Escher B, Leusch F. 2012a. *Bioanalytical tools in water quality assessment*. IWA
485 Publishing, London, UK.
486 Escher BI, Ait-Aissa S, Behnisch PA, Brack W, Brion F, Brouwer A, Buchinger S,
487 Crawford S, Hamers THM, Hettwer K, Hilscherova K, Hollert H, Kase R,
488 Kienle C, Legradi J, Tuerk J, van der Oost R, Vermeirssen E, Neale PA. 2018.
489 Effect-based trigger values for *in vitro* and *in vivo* bioassays performed on
490 surface water extracts supporting the environmental quality standards (EQS)
491 of the European Water Framework Directive. *Sci Tot Environ* 628-629:748-
492 765.
493 Escher BI, Allinson M, Altenburger R, Bain P, Balaguer P, Busch W, Crago J,
494 Humpage A, Denslow ND, Dopp E, Hilscherova K, Kumar A, Grimaldi M,
495 Jayasinghe BS, Jarosova B, Jia A, Makarov S, Maruya KA, Medvedev A,
496 Mehinto AC, Mendez JE, Poulsen A, Prochazka E, Richard J, Schifferli A,
497 Schlenk D, Scholz S, Shiraishi F, Snyder S, Su G, Tang J, van der Burg B,
498 van der Linden S, Werner I, Westerheide SD, Wong CKC, Yang M, Yeung B,
499 Zhang X, Leusch FDL. 2014. Benchmarking organic micropollutants in
500 wastewater, recycled water and drinking water with in vitro bioassays. *Environ*
501 *Sci Technol* 48:1940-1956.
502 Escher BI, Dutt M, Maylin E, Tang JYM, Toze S, Wolf CR, Lang M. 2012b. Water
503 quality assessment using the AREc32 reporter gene assay indicative of the
504 oxidative stress response pathway. *J Environ Monitor* 14:2877-2885.
505 Escher BI, Neale PA, Leusch FDL. 2015. Effect-based trigger values for in vitro
506 bioassays: Reading across from existing water quality guideline values. *Water*
507 *Res* 81:137-148.
508 Escher BI, van Daele C, Dutt M, Tang JYM, Altenburger R. 2013. Most oxidative
509 stress response in water samples comes from unknown chemicals: the need
510 for effect-based water quality trigger values. *Environ Sci Technol* 47:7002-
511 7011.

512 Finney DJ. 1951. The statistical analysis of slope-ratio assays. *Journal of General*
513 *Microbiology* 5:223-230.
514 Fox DR, Landis WG. 2016. Don't be fooled-a no-observed-effect concentration is no
515 substitute for a poor concentration-response experiment. *Environ Toxicol*
516 *Chem* 35:2141-2148.
517 Green JW, Springer TA, Staveley JP. 2013. The drive to ban the NOEC/LOEC in
518 favor of ECx is misguided and misinformed. *Int Environ Ass Manag* 9:12-16.
519 Hashmi MAK, Escher BI, Krauss M, Teodorovic I, Brack W. 2018. Effect-directed
520 analysis (EDA) of Danube River water sample receiving untreated municipal
521 wastewater from Novi Sad, Serbia. *Sci Tot Environ* 624:1072 - 1081.
522 Hebert A, Felier C, Lecarpentier C, Neale P, Schlichting R, Thibert S, Escher B.
523 2018. Bioanalytical assessment of adaptive stress responses in drinking water
524 as a tool to differentiate between micropollutants and disinfection by-products.
525 *Water Res* 132:340-349.
526 Howard GJ, Schlezinger JJ, Hahn ME, Webster TF. 2010. Generalized
527 Concentration Addition Predicts Joint Effects of Aryl Hydrocarbon Receptor
528 Agonists with Partial Agonists and Competitive Antagonists. *Environ Health*
529 *Perspect* 118:666-672.
530 Howard GJ, Webster TF. 2009. Generalized concentration addition: A method for
531 examining mixtures containing partial agonists. *Journal of Theoretical Biology*
532 259:469-477.
533 ISO13829:2000. 2000. Water quality – Determination of the genotoxicity of water and
534 waste water using the umu-test. International Organization for Standardization
535 (ISO), Geneva, Switzerland.
536 Judson R, Houck K, Martin M, Richard AM, Knudsen TB, Shah I, Little S, Wambaugh
537 J, Setzer RW, Kothya P, Phuong J, Filer D, Smith D, Reif D, Rotroff D,
538 Kleinstreuer N, Sipes N, Xia MH, Huang RL, Crofton K, Thomas RS. 2016.
539 Analysis of the Effects of Cell Stress and Cytotoxicity on In Vitro Assay Activity
540 Across a Diverse Chemical and Assay Space. *Tox Sci* 152:323-339.
541 König M, Escher BI, Neale PA, Krauss M, Hilscherová K, Novák J, Teodorović I,
542 Schulze T, Seidensticker S, Kamal Hashmi MA, Ahlheim J, Brack W. 2017.
543 Impact of untreated wastewater on a major European river evaluated with a
544 combination of *in vitro* bioassays and chemical analysis. *Environ Pollut*
545 220:1220-1230.

546 Limpert E, Stahel WA, Abbt M. 2001. Log-normal distributions across the sciences:
547 Keys and clues. *Bioscience* 51:341-352.
548 Neale PA, Ait-Aissa S, Brack W, Creusot N, Denison MS, Deutschmann B,
549 Hilscherová K, Hollert H, Krauss M, Novák J, Schulze T, Seiler T-B, Serra H,
550 Shao Y, Escher BI. 2015. Linking in Vitro Effects and Detected Organic
551 Micropollutants in Surface Water Using Mixture-Toxicity Modeling. *Environ Sci*
552 *Technol* 49:14614-14624.
553 Neale PA, Altenburger R, Ait-Aissa S, Brion F, Busch W, de Aragão Umbuzeiro G,
554 Denison MS, Du Pasquier D, Hilscherova K, Hollert H, Morales DA, Novac J,
555 Schlichting R, Seiler T-B, Serra H, Shao Y, Tindall AJ, Tollefsen KE, Williams
556 TD, Escher BI. 2017a. Development of a bioanalytical test battery for water
557 quality monitoring: Fingerprinting identified micropollutants and their
558 contribution to effects in surface water. *Water Res* 123:734-750.
559 Neale PA, Brack W, Ait-Aissa S, Busch W, Hollender J, Krauss M, Maillot-Maréchal
560 E, Munz NA, Schlichting R, Schulze T, Vogler B, Escher B. 2018. Solid-phase
561 extraction as sample preparation of water samples for cell-based and other *in*
562 *vitro* bioassays. *Environ Sci Process Impacts* 20:493-504.
563 Neale PA, Munz NA, Ait-Aissa S, Altenburger R, Brion F, Busch W, Escher BI,
564 Hilscherova K, Kienle C, Novak J, Seiler T-B, Shao Y, Stamm C, Hollender J.
565 2017b. Integrating Chemical Analysis and Bioanalysis to Evaluate the
566 Contribution of Wastewater Effluent on the Micropollutant Burden in Small
567 Streams. *Sci Tot Environ* 576:785-795.
568 Nivala J, Neale PA, Haasis T, Kahl S, König M, Müller R, Reemtsma T, Schlichting
569 R, Escher B. 2018. Application of bioanalytical tools to evaluate treatment
570 efficacy of conventional and intensified treatment wetlands. *Environ Sci: Water*
571 *Res Technol* 4:206-217.
572 Ritz C. 2010. Toward a unified approach to dose-response modeling in
573 ecotoxicology. *Environ Toxicol Chem* 29:220-229.
574 Ritz C, Cedergreen N, Jensen JE, Streibig JC. 2006. Relative potency in nonsimilar
575 dose-response curves. *Weed Sci* 54:407-412.
576 Schmitt S, Reifferscheid G, Claus E, Schlusener M, Buchinger S. 2012. Effect
577 directed analysis and mixture effects of estrogenic compounds in a sediment
578 of the river Elbe. *Environ Sci Pollut Res* 19:3350-3361.

579 Scholze M, Boedeker W, Faust M, Backhaus T, Altenburger R, Grimme LH. 2001. A
580 general best-fit method for concentration-response curves and the estimation
581 of low-effect concentrations. *Environ Toxicol Chem* 20:448-457.
582 Shukla SJ, Huang RL, Austin CP, Xia MH. 2010. The future of toxicity testing: a focus
583 on *in vitro* methods using a quantitative high-throughput screening platform.
584 *Drug Discov Today* 15:997-1007.
585 Slob W. 2002. Dose-response modeling of continuous endpoints. *Tox Sci* 66:298-
586 312.
587 Tang JYM, Busetti F, Charrois JWA, Escher BI. 2014. Which chemicals drive
588 biological effects in wastewater and recycled water? *Water Res* 60:289-299.
589 van der Oost R, Sileno G, Suarez-Munoz M, Nguyen MT, Besselink H, Brouwer A.
590 2017. SIMONI (smart integrated monitoring) as a novel bioanalytical strategy
591 for water quality assessment: Part I–model design and effect-based trigger
592 values. *Environ Toxicol Chem* 36:2385-2399.
593 Villeneuve DL, Blankenship AL, Giesy JP. 2000. Derivation and application of relative
594 potency estimates based on *in vitro* bioassay results. *Environ Toxicol Chem*
595 19:2835-2843.
596 Volund A. 1978. Application of 4-parameter logistic model to bioassay - comparison
597 with slope ratio and parallel line models. *Biometrics* 34:357-365.
598 Wagner M, Vermeirssen ELM, Buchinger S, Behr M, Magdeburg A, Oehlmann J.
599 2013. Deriving bio-equivalents from in vitro bioassays: Assessment of existing
600 uncertainties and strategies to improve accuracy and reporting. *Environ*
601 *Toxicol Chem* 32:1906-1917.
602 Windal I, Denison MS, Birnbaum LS, Van Wouwe N, Baeyens W, Goeyens L. 2005.
603 Chemically activated luciferase gene expression (CALUX) cell bioassay
604 analysis for the estimation of dioxin-like activity: Critical parameters of the
605 CALUX procedure that impact assay results. *Environ Sci Technol* 39:7357-
606 7364.
607