Griffith UNIVERSITY

smartwater Research Centre science securing our future

Analysis of bioanalytical data using non-linear regression in Excel

Frederic DL Leusch and Erik Prochazka

Smart Water Research Centre & School of Environment, Griffith University, Parklands Drive, Gold Coast, Qld 4222, Australia

Introduction

The value of *in vitro* bioassay testing in environmental monitoring applications has over the past decade become clear to researchers and regulators (1). In particular, *in vitro* bioassays can help overcome some of the common limitations of chemical analysis, such as detection of non-target compounds, transformation products and mixture effects. There is still, however, some concern about the reliability and robustness of *in vitro* bioassays compared to well-established chemical analysis methods. One issue of particular concern is the perceived inter-laboratory variability with bioassay results.

During a previous inter-laboratory research project (2), adoption of a standardized bioassay data analysis method greatly reduced initial inter-laboratory differences, and standardization may therefore also noticeably reduce reported differences between laboratories and research groups.

The following poster focuses on bioanalytical techniques (*i.e.*, *in vitro* bioassay) in the context of water quality assessment, although some of the concepts discussed are also valid for *in vivo* bioassays.

Background to bioassay data analysis

When testing a sample for *in vitro* biological activity, it is important to test a serial dilution of that sample over several orders of magnitude, from most to least concentrated. This produces a full concentration-effect curve rather than a single value. The resulting data is presented as effect (generally expressed as a fraction of the maximal effect with the standard compound) *vs.* concentration (for model compounds) or relative enrichment factor (for water samples) on a logarithmic scale. The full concentration-effect curve provides a significantly more robust representation of biological activity than a single concentration data would, and ensures that potential masking effects (*e.g.*, cytotoxicity interference) are detected (Fig. 1).

Once a concentration-effect plot is produced, the data usually fits a sigmoid function (Fig. 2), although in some instance the top of the curve is not reached (in the case of cytotoxicity interference for example; Fig. 1). There are then four common methods to express a single quantitative biological activity from the concentration-effect curve, each with its advantages and limitations (Table 1).

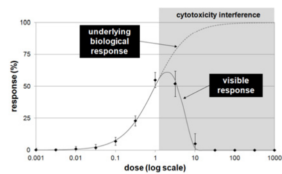


Fig. 1. Typical bioassay response showing the masking effect due to cytotoxicity. Adapted from Fig. 7.3 in (1)

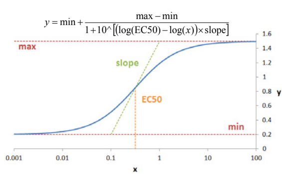


Fig. 2. Sigmoid curve

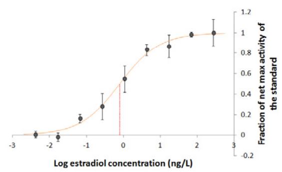


Fig. 3. Sample bioassay data: Estradiol standard curve in the E-SCREEN assay

Table 1. Typical methods to summarize a concentration-effect curve into a single number

Name	Graph	Method	Advantages	Limitations	Notes
First significant difference (FSD)		ANOVA followed by Tukey's post-hoc is used to determine the first data point in the series that is significantly different from the baseline.	- Simple method based on proven statistical methods. - Does not require curve-fit, only ANOVA. - Does not require a full effect curve.	- A single point is used. - Some assays have very low variability in the baseline, and even a small response may be statistically significant eventhough it may not be biologically significant.	The first significant difference method, while statistically sound, is not always relevant as it can vary significantly between assays depending on the variability around the baseline.
Interpolation in the linear range		The lowest point above EC_{LOQ} but below EC_{50} (i.e., within the linear range of the sigmoid curve) with a coefficient of variation of <15% is used.	- Simple method. - Does not require curve-fit. - Easily automated in Excel.	- A single point is used. - Requires relatively narrow dilution steps (2-4×) to ensure that at least one point will be within the desired EC range.	The interpolation in the linear range method provides a good compromise between biological relevance and practical limitations, but it uses only one datum from the whole concentration-effect curve.
EC_{50}		A sigmoid curve is fitted to the data and the EC_{50} is derived from the curve-fit.	- Incorporates all available data points, thereby reducing the effect of variability from individual concentrations. - Using the inflexion point of the sigmoid curve (EC_{50}) means that this method is less impacted by minor differences in slope.	- Requires a full dose response (from 0 to 100%), which is not always possible to generate when dealing with samples with low biological activity.	Even with pre-concentration, it is thankfully rare to find environmental samples that produce a full effect in an *in vitro* bioassay, even at the highest concentration tested. This means that the EC_{50} method, while robust from a purely biological point of view, is generally not possible.
Trigger level		A sigmoid curve is fitted to the data and the concentration (or relative enrichment factor) required to reach the trigger level (usually EC_{LOQ}) is used.	- Simple method. - Does not require a full effect curve. - Incorporates all available data points, thereby reducing the effect of variability from individual concentrations. - Easily automated in Excel.	- The method to set the trigger level can vary between different assays (*e.g.*, LOQ, LOD, specific induction ratio).	The trigger level method is our method of choice for bioanalytical data analysis: it is practical, applicable to samples with low biological activity, easy to automate and implement in Excel and uses the entire dataset (by using the line of best fit) rather than a single datum.

Automation of bioassay data analysis in Excel

Excel is a component of the Microsoft Office suite and is easily installed on most modern computers. The trigger level method for bioassay data analysis can easily be implemented and automated in Excel using the Solver Add-in to produce the curve of best fit using a least-square regression method, as described below. The Excel file itself can be downloaded from http://fredleusch.swifthost.net/research/bda, or by scanning the QR code at the bottom right of this poster.

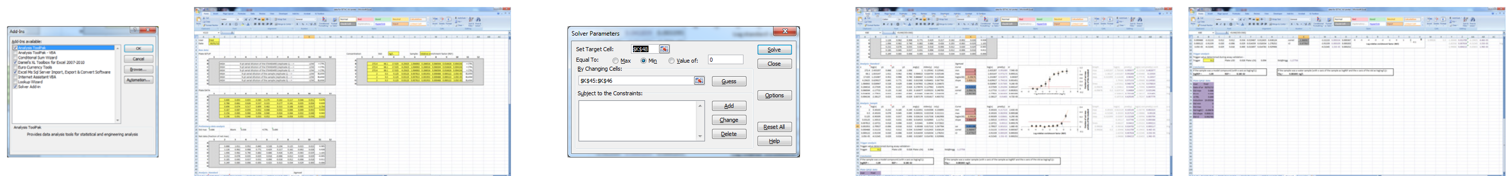
Step 1: Install the Solver Add-in: Click on the `Office ribbon > Excel Options > Add-Ins > Manage: Excel Add-Ins > Go`, and tick the appropriate check box.

Step 2: Enter date, username, plate details (*i.e.*, units, concentrations) and paste the raw data – all of the yellow cells need to be edited. Your assay plates must be set up as depicted in "Plate SETUP".

Step 3: Scroll down the sheet to "Analysis: Standard", click on the blue cell labelled "ssr" (`K48`), then go to `Data > Solver`. Enter the parameters as in the screenshot below and click `Solve`. This will let solver modify the logEC50 (`K45`) and slope (`K46`) to make the sum of squared residuals (`K48`) as small as possible – the principle of least-square regression.

Step 4: Repeat the Solver process for your sample data: set the blue "ssr" cell (`K60`) to min by changing only the logEC50 (`K57`) this time (it is better if the slopes of the standard curve and the sample are the same to ensure parallelism).

Step 5: Enter the trigger EC value (`B67`), previously determined during assay validation (generally EC_{LOQ} calculated as 10× standard deviation of the baseline). Scroll down the sheet to "Conclusion" to see the potency, as logREP for model compounds or toxic equivalent concentration for water samples.



Conclusions

Using this simple approach in the widely available Excel program, *in vitro* bioassay data analysis can be standardized. Applying a standardized approach to data analysis may help reduce the inter-laboratory variability, hopefully removing one problem that has negatively impacted the acceptance of bioanalytical tools by the wider scientific and regulatory community.

Acknowledgment

The authors wish to thank Prof Beate Escher and Prof Jochen Mueller for stimulating discussion during the development of this approach and Dr Maud Achard for editorial comments.

References

(1) Escher and Leusch (2012). Bioanalytical tools in water quality assessment. With contributions by Chapman and Poulsen. IWA Publishing; **(2)** Leusch *et al.* (2010). Environ Sci Technol 44: 3853-3860.

http://fredleusch.swifthost.net/research/bda