BioRssay: An R Package for Analyses of Bioassays and Probit Graphs

Piyal Karunarathne
  Uppsala University: Uppsala Universitet

Nicolas Pocquet
  Pasteur Institute: Institut Pasteur

Pierrick Labbé
  Montpellier University: Universite de Montpellier

Pascal Milesi (✉ pascal.milesi@scilifelab.uu.se)
  Uppsala University   https://orcid.org/0000-0001-8580-4291

Short report

Abstract

Dose-response relationships reflect the effects of a substance on organisms, and are widely used in broad research areas, from medicine and physiology, to vector control and pest management in agronomy. Furthermore, reporting on the response of organisms to stressors is an essential component of many public policies (e.g. public health, environment), and assessment of xenobiotic responses is an integral part of the World Health Organization recommendations. Building upon an R script that we previously made available, and considering its popularity, we have now developed a software package in the R environment, BioRssay, to efficiently analyze dose-response relationships. It has more user-friendly functions, more flexibility, and proposes an easy interpretation of the results. The functions in the BioRssay package are built on robust statistical analyses to compare the dose/exposure-response of various bioassays and effectively visualize them in probit-graphs.

Background

Bioassays aim at evaluating the potency of a compound. It usually consists in measuring the response of a subject (e.g. organisms, populations, tissues, cells lines) to increasing doses (or intensities or exposure times) of a stimulus (most of the time a xenobiotic or a chemical) to quantify specific dose-response relationships (also known as exposure-response relationships). Because of their effectiveness, dose-response relationship analyses are widely used in a large spectrum of scientific disciplines (e.g. epidemiology, microbiology, toxicology, environment quality monitoring, vector and pest control, and parasite biology). A prime example of such analysis is the monitoring of resistance to xenobiotics. Since the 1950s, xenobiotics (e.g., insecti-, pesti-, fungicides) have been widely used to control populations of vectors or pests [1]. However, as a counter result, resistance mechanisms to such substances have been selected in targeted populations, undermining their efficiency [2]. Therefore, establishing and comparing the resistance levels of various populations to various xenobiotics is at the core of the World Health Organization's (WHO) recommendations in order to define/adjust vector control strategies. It is usually done by exposing batches of individuals (adults or larvae) to varying doses of the xenobiotic to assess their responses (e.g. mortality or knock-down effect).

Despite its use in many fields, there has been a substantial lack of easily accessible statistical infrastructure for the analysis of dose-response relationships. In 2013, we developed an R-script with a robust statistical background to describe and compare the dose-mortality relationship [3], and it has since gained much popularity and is widely used for similar studies e.g. [4–10]. In order to make it more user-friendly and easily accessible to the science community, we have now developed it into an R package called 'BioRssay', with more flexibility and an improved presentation of the results.

Workflow

BioRssay is a comprehensive compilation of scripts in R language [11] designed to analyze dose-response relationships (or exposure-response: mortality, knock-down effect, etc.) from bioassays of one

or more subjects (populations, strains, lines, tissues, cells, etc.). This package provides a complete analytic workflow from data quality assessment to statistical analyses and data visualization (see Fig. 1).

a. In the first step, the base mortality in the controls (*i.e.* without the stimulus to be tested) is taken into account to control for the mortality (or effect) associated with the experiment itself, regardless of the exposure. Data are adjusted following Abbott's formula using a maximum likelihood approach (*optim* function, method "L-BFGS-B", *stats* R package[11]), to estimate correction factors (Eq. 1), while taking into account heterogeneity in the mortality rates between replicates (see example in Table 1) [12]:

$$1 - \frac{m_T}{m_C} \text{ Eq. 1,}$$

where m_T and m_C are the survival rates in the treatments and the controls, respectively. By default, the correction is applied if mortality in the control is higher than 5%, but any threshold can be specified.

b. Data are then probit-transformed and analyzed using a generalized linear model (probit-link function). The coefficients of the linear regressions (*i.e.* slope and intercept, with their standard errors) are reported (see example in Table 2). These coefficients are used to test the linearity of the log-dose response for each subject using a chi-square test of homogeneity between model predictions and probit-transformed data (Eq. 2); significant deviations from linearity may for example reflect mixed populations, or a threshold effect.

$$\chi^2 = \sum \frac{(observed - predicted)^2}{predicted} \text{ Eq. 2,}$$

c. By default, lethal doses, LD (aka lethal concentrations, LC), are computed at 25%, 50%, and 95% of mortality with their corresponding 95% confidence intervals (CI), following the approach developed by Johnson et al. [10] from Finney [11], which allows taking data heterogeneity into account (Table 1); otherwise, any level of LD and CI can be specified by the user (Table 2). Resistance ratio (RR) with 95% confidence intervals are calculated according to Robertson and Preisler [13]. They measure the magnitude of the difference in dose-response of two subjects when exposed to the same stimulus. For a given LD (LD25, 50, and 95 by default), the LD of a focal subject is divided by the LD of the subject with the lowest one (usually a susceptible reference, Table 2).

d. When different subjects (population, strains, lines, etc.) are exposed to the same stimulus, we offer the possibility to test whether they present significant differences in their responses. A likelihood ratio test (LRT) is conducted to compare a null model assuming that all the data come from the same subject, where the dose is the only explanatory variable (Eq. 3), with a model considering the dose nested in the subject effect (Eq. 4).

$$y_i = a_i + \epsilon_i \text{ Eq. 3, and } y_{ij} = \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ij} \text{ Eq. 4,}$$

where y is the mortality rate measured in sample i from subject j, exposed to a dose α, and where β is the subject-specific fixed-effect, γ is the interaction between α (dose) and β (subject) effects and ϵ is the error term following a quasi-binomial distribution. If the two models are significantly different (*i.e.* at least two subjects show differing responses either in magnitude of the response and / or slope) and that there are more than two subjects, pairwise comparisons (LRT) are implemented to test for statistical significance of the difference between all pairs of subjects; the Holm-Bonferroni method is used to control for the family-wise error inherent to multiple testing [14].

Note that if at least one of the subjects failed the linearity test described above (point b.), the validity of comparing the slope of the dose-mortality response is, at best, highly questionable (not recommended). If many subjects are present and only one (few) fails the linearity tests, we do recommend removing this (these) specific subject(s) before conducting the test.

e. Customizable (e.g. colors, symbols, title) plots of the probit-transformed regressions are drawn (Fig. 2A and B). By default, if the probit-transformed response significantly deviates from linear regression, the data are connected by segments and the regression and associated CI are not plotted (e.g. Figure 2B, *Acerkis*, red, and *AgRR5*, green, strains). Confidence intervals around the regressions can be removed or added, and users can specify any levels of CI.

Package accessibility and concluding remark

The package is available freely on GitHub at https://github.com/milesilab/BioRssay (https://doi.org/10.5281/zenodo.5172072). We also provide a comprehensive workflow and tutorials, from data preparation to the interpretation of the results, in a dedicated GitHub page https://milesilab.github.io/BioRssay/. Further, the package carries example data sets for self-tests and more data can be downloaded at https://github.com/milesilab/DATA. The code is maintained by Piyal Karunarathne (piyalkarumail@yahoo.com). For suggestions or further development please contact the corresponding authors, P. Labbé and P. Milesi.

The BioRssay package can be installed in the R environment using the following code:

```
if (!requireNamespace("devtools", quietly = TRUE))
```

```
install.packages("devtools")
```

```
devtools::install_github("milesilab/BioRssay")
```

Abbreviations

LD – Lethal dose

WHO – World health organization

CI – Confidence interval

LRT – Likelihood ratio tests

RR – Resistance ratio

Declarations

- Ethics approval and consent to participate: **Not applicable**
- Consent for publication: **Not applicable**
- Availability of data and materials: accessible at https://github.com/milesilab/DATA
- Competing interests: **Not applicable**
- Funding: **Not applicable**
- Authors' contributions: PM and PL compiled the methodology and statistical approach; PL, PM, PN, and PK wrote the scripts and PK compiled the functions and the software package. PK, PL, and PM prepared the manuscript.
- Acknowledgments: We want to thank Jérôme Chopard, Haoues Alout, Mylène Weill, and Nicole Pasteur for their valuable comments on earlier versions of the script and its outputs.

References

1. Carson R. Silent Spring. Boston: Houghton Mifflin; 1962.
2. Labbé P, David J-P, Alout H, Milesi P, Djogbénou L, Pasteur N, et al. 14 - Evolution of Resistance to Insecticide in Disease Vectors. In: Tibayrenc MBT-G and E of ID (Second E, editor. London: Elsevier; 2017. p. 313–39.
3. Milesi P, Pocquet N, Labbé P. BioRssay: a R script for bioassay analyses. BioRssay a R Script. bioassay 2013.
4. Alout H, Labbé P, Berthomieu A, Makoundou P, Fort P, Pasteur N, et al. High chlorpyrifos resistance in *Culex pipiens* mosquitoes: strong synergy between resistance genes. Heredity; 2016;116:224–31.
5. Pocquet N, Darriet F, Zumbo B, Milesi P, Thiria J, Bernard V, et al. Insecticide resistance in disease vectors from Mayotte: an opportunity for integrated vector management. Parasit Vectors; 2014;7:299.
6. Badolo A, Sombié A, Pignatelli PM, Sanon A, Yaméogo F, Wangrawa DW, et al. Insecticide resistance levels and mechanisms in Aedes aegypti populations in and around Ouagadougou, Burkina Faso. 2019. PLoS Negl Trop Dis.019;13:e0007439.
7. Assogba BS, Milesi P, Djogbénou LS, Berthomieu A, Makoundou P, Baba-Moussa LS, et al. The ace-1 locus is amplified in all resistant Anopheles gambiae mosquitoes: fitness consequences of homogeneous and heterogeneous duplications. PLoS Biol. 14:e2000618.
8. Yaméogo F, Wangrawa DW, Sombié A, Sanon A, Badolo A. Insecticidal activity of essential oils from six aromatic plants against Aedes aegypti, dengue vector from two localities of Ouagadougou,

Burkina Faso. Arthropod Plant Interact. 2021;1–8.

9. Epelboin Y, Wang L, Giai Gianetto Q, Choumet V, Gaborit P, Issaly J, et al. CYP450 core involvement in multiple resistance strains of Aedes aegypti from French Guiana highlighted by proteomics, molecular and biochemical studies. PLoS One. 16:e0243992.

10. Perrier S, Moreau E, Deshayes C, El-Adouzi M, Goven D, Chandre F, et al. Compensatory mechanisms in resistant Anopheles gambiae AcerKis and KdrKis neurons modulate insecticide-based mosquito control. Commun Biol. 2021;4:1–16.

11. R Core Team. A language and environment for statistical computing, R Foundation for Statistical Computing; 2013. 2020.

12. Abbott WS. A method of computing the effectiveness of an insecticide. J Econ Entomol. 1925;18:265–7.

13. Robertson JL, Preisler HK. No Title. Pestic. bioassays with arthropods. CRC Boca Raton, FL; 1992.

14. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat. 1979;65–70.

Tables

Table 1

A subset of bioassays conducted on three strains of *A. gambiae* mosquitoes exposed to increasing doses of Temephos® insecticide (data from Pocquet et al. 2014.; assay2 data set in the example data in the BioRssay package). The first four columns are the mandatory input format. *Mort* and *Probmort* are Abbott's corrected mortalities and probit-transformed mortalities, respectively.

Insecticide	Strain	Dose	Total	Dead	Replicate	Date	Mort	Probmort
temephos	KIS-ref	0.002	97	46.74	1	2026-01-11	0.481	-0.045
temephos	KIS-ref	0.003	96	67.85	1	2026-01-11	0.706	0.544
temephos	KIS-ref	0.004	98	88.953	1	2026-01-11	0.907	1.326
temephos	DZOU	0.001	97	4.04	1	2004-08-10	0.041	-1.731
temephos	DZOU	0.002	97	20.20	1	2004-08-10	0.208	-0.812
temephos	DZOU	0.004	100	31.29	1	2004-08-10	0.312	-0.487
temephos	DZOU	0.007	95	51.55	1	2004-08-10	0.542	0.107
temephos	DZOU2	0.002	97	24.25	1	2004-08-10	0.250	-0.674
temephos	DZOU2	0.004	100	35.33	1	2004-08-10	0.353	-0.376
temephos	DZOU2	0.007	95	55.59	1	2004-08-10	0.585	0.215
temephos	DZOU2	0.010	97	68.70	1	2004-08-10	0.708	0.5485

Table 2
Parameters estimated from the probit-transformed data (data available at https://github.com/milesilab/DATA/blob/main/BioAssays/Test.BioRssay.txt).

Strains	Linear regression parameters			Lethal Dose 50 % (LD$_{50}$)			Resistance Ratio (RR$_{50}$)		
	Slope (± SE)	Intercept (± SE)	Chi(p)[a]	LD	LD$_{min}$	LD$_{max}$	RR	RR$_{min}$	RR$_{max}$
Acerkis	5.88 ± 0.37	-10.30 ± 0.67	1.00	57	53	60	207	103	418
AgRR5	8.52 ± 0.70	-16.17 ± 1.36	0.96	79	75	83	290	144	583
Kisumu[b]	7.69 ± 0.61	4.34 ± 0.33	1.00	0.27	0.26	0.29	-	-	-

[a] Chi(p) is the *p*-value of the chi-square test of homogeneity

[b] Kisumu strain has the lowest LD$_{50}$ and is used as the reference to compute the resistance ratios (RR$_{50}$)
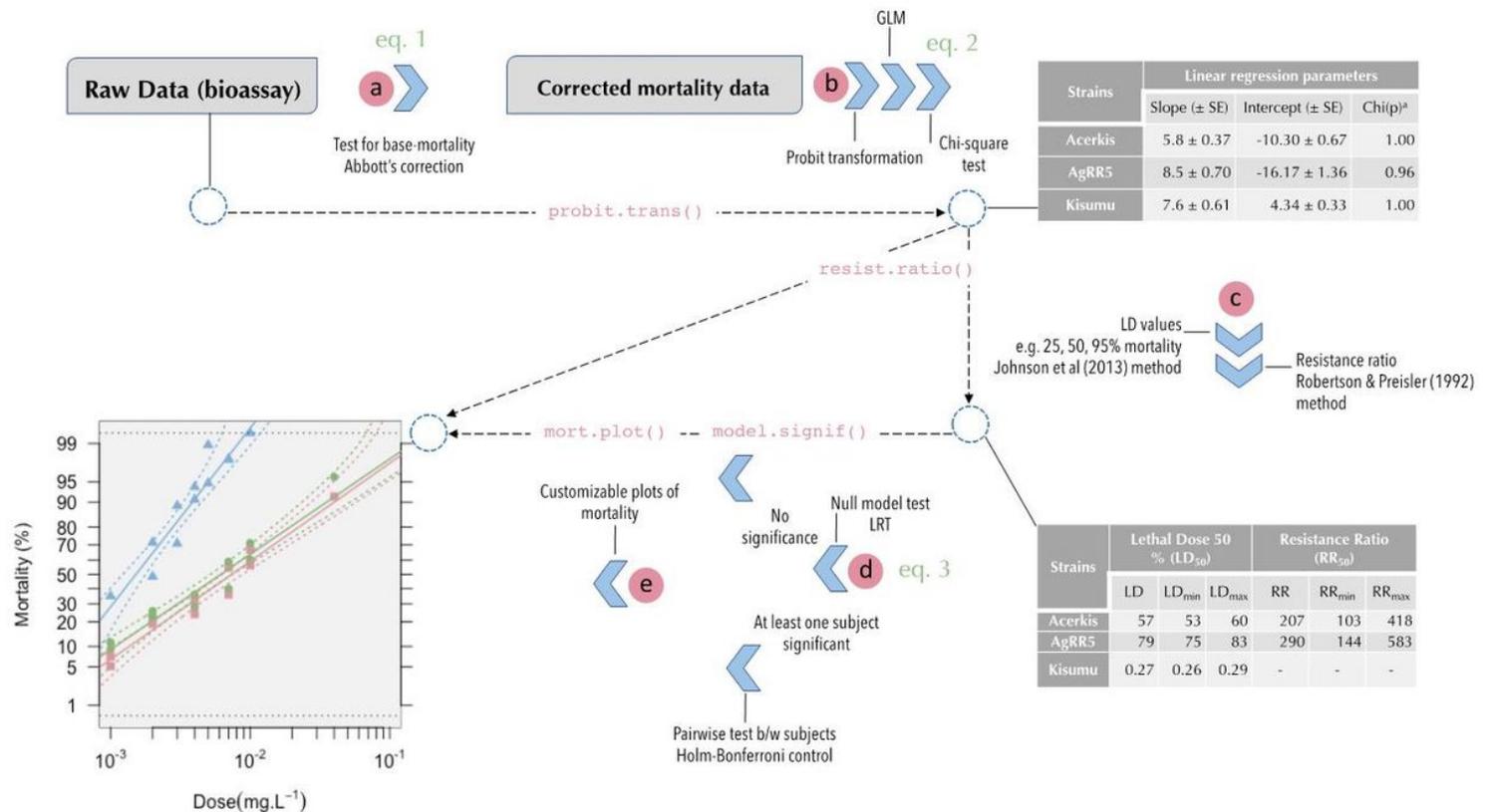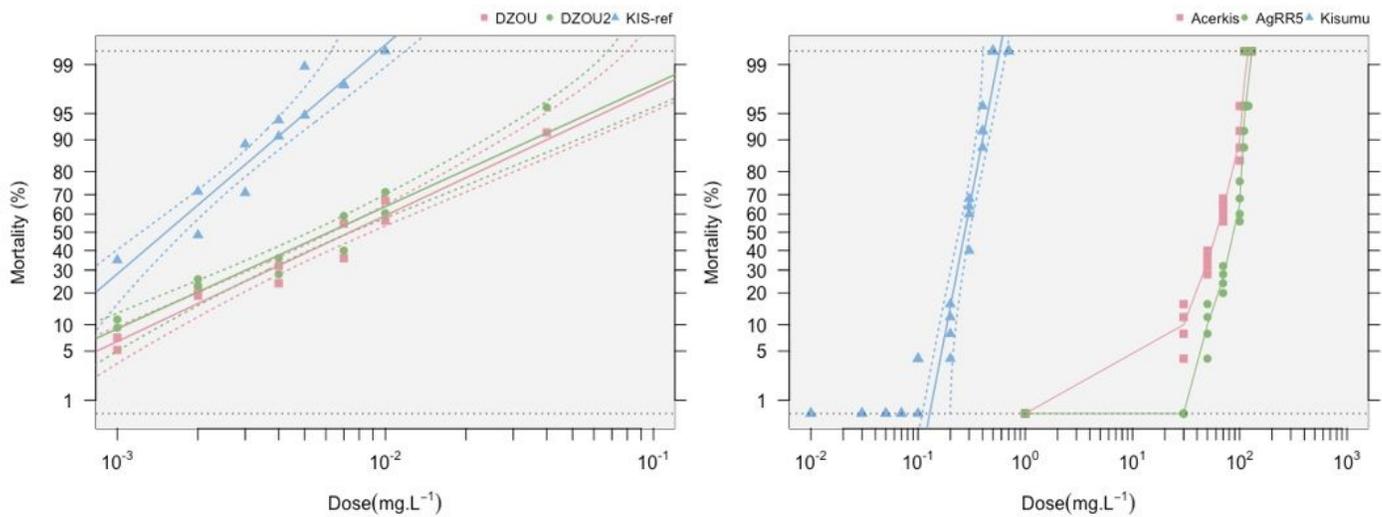
Figures

Figure 1

General workflow of the BioRssay package. a. to e. refer to the main steps of the workflow fully detailed in the main text.

Figure 2

Probit graphs generated by the BioRssays R package. A. Linear relationships between probit-transformed mortality rates and log-dose for different mosquito's strains. Kis-ref (blue triangles) is the susceptible reference strain. DZOU (red square) and DZOU2 (green dots) show similar resistance levels, significantly higher than the reference strain (Tab. 2). B. Same as A., but for Acerkis (red squares) and AgRR5 (green dots) the relation is not linear and dots are connected by segments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Picture1.png