



CalIBRi: a web interface to calculate Integrated Biomarker Index

Simon Devin, Pierre-Yves Arnould, Laetitia Minguez, Sandrine Pain-Devin,
Virginie Jouffret

► To cite this version:

Simon Devin, Pierre-Yves Arnould, Laetitia Minguez, Sandrine Pain-Devin, Virginie Jouffret. Cal-IBRi: a web interface to calculate Integrated Biomarker Index. Environmental Science and Pollution Research, 2023, 10.1007/s11356-023-27447-7 . hal-04093993

HAL Id: hal-04093993

<https://hal.univ-lorraine.fr/hal-04093993>

Submitted on 10 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NoDerivatives 4.0 International License

CalIBRi: a web interface to calculate Integrated Biomarker Index

Simon Devin^{1‡}, Pierre-Yves Arnould², Laetitia Minguez¹, Sandrine Pain-Devin¹, Virginie Jouffret^{1,3}

1: Université de Lorraine, CNRS, LIEC, F-57000 Metz, France

2: UMS OTELo, CNRS, F-54000, Nancy

3: BigA Core Facility, Centre de Biologie Intégrative, Université de Toulouse, F-31062 Toulouse, France

‡: corresponding author. Email: simon.devin@univ-lorraine.fr

Preamble

This is an addendum rather than an erratum.

To overcome the challenge raised by multibiomarker approaches in ecotoxicology (analyze and interpret the datasets in a synthetic and integrative manner), we proposed in 2014 an integrative index, the Integrated Biomarker Response - IBR (Devin et al. 2014).

This index was quite successful (129 citations in Web of Sciences, 169 in Google Scholar), but its usage remains limited by the need to use a quite complex R code, and lack of solution to produce a robust statistical analysis of the data produced. We thus developed a web interface to facilitate the appropriation of the index by researchers and stakeholders. It was built for up to 9 biomarkers, as described in the original paper, and for up to 16 conditions.

Description of the interface

The interface is accessible through the following link: <https://shiny.otelo.univ-lorraine.fr/calibri/R/>.

It is structured in 3 main different sheets and a help page, and provides sample datasets that can be used to (1) test the interface and (2) get an example of the data file organization. The first page is dedicated to file loading. Datasets are checked to fit with the format used by the interface. This first pages also allows for biomarker selection. This page proposes a visual checking of the first lines of the uploaded tables.

The different steps of the IBR calculation are then performed, and the second page contains all the results related to this calculation: Standardized values, histogram of the IBR values calculated through permutations, boxplot of IBR values in the different experimental conditions and several versions of the radar plot (points only, points and surfaces in black and white or in color). Numeric results can be downloaded in a .csv format, and graphics in .png or .pdf format.

Finally, a last page is dedicated to the statistical analysis of differences between experimental conditions.

A detailed help page gathers all the necessary information to understand IBR calculation, permutation procedure, website utilization and statistical approach developed.

Novelty compared to the initial paper

(1) The number of IBR obtained through permutation was not $(k-1)!$, but $(k-1)!/2$: indeed, we miss to integrate in the original paper that the same permutation can be obtained in two different way, considering you look at the radar diagram in clockwise or counter-clockwise rotation.

(2) a statistical module was added. You can compare IBR values between categories. Values are always normally distributed and homoscedasticity are respected, allowing using paired two sample t-test to assess differences between conditions. As IBR procedure results in a huge sample size of IBR values, there is a bias in statistical tests that become too powerful. As a consequence, statistically significant differences are systematically found, but that not always make sense from a biological point of view. To get around this problem, we first set the magnitude of the effect size (a dimensionless parameter that quantifies IBR difference between two conditions) to 1, which is basically meaningful. With this basis, we then calculated the sample size allowing such differences, by setting the statistical power to 0.8, which is a value generally considered as good. This led to a sample size equal by default to the geometric mean of the value number in each biomarker/ category condition. CalIBRi thus propose by default this value as the random sample size, but it can be changed by the users.

To encompass the variability of IBR among permutations, we repeated by default 100 times this random sampling of IBR values among all the ones calculated within each category, and performed each time a bilateral paired student t-test between the two categories. The p-value of each test was then corrected for multiple comparisons following the procedure of Benjamini and Hochberg (1995). Finally, the p-value we kept was the mean of the corrected p-values.

Through check box you can select couple of category that you wish test statistically. As the number of tests influences the result, be careful and choose a set of important category pair and avoid to multiply the analysis. When all pairs are selected, click on *Run* button. A resume dataframe lists all paired two samples t-test computed and the mean of adjusted p-value.

References

Benjamini, Y. and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B*, **57**, 289-300.

Devin, S., Burgeot, T., Giambérini, L., Minguez L. and Pain-Devin S. (2014) The integrated biomarker response revisited: optimization to avoid misuse. *Environ Sci Pollut Res* **21**, 2448–2454.