

metadata

Gustaf M.O. Ekelund Ugge

2022-09-06

This is an overview of the code and data used in ‘Transcriptional responses as biomarkers of general toxicity: a systematic review and meta-analysis on metal-exposed bivalves’.

Dataset

The ‘dataset.xlsx’ file contains the dataset used for the meta-analysis. Each column of the dataset is briefly explained below.

study

Study ID; unique to each publication and assigned during the full-text selection procedure.

reference

Short reference (first author and year) to each publication.

species

Latin name of the study species.

habitat

Habitat type of the study species; marine or freshwater.

tissue

The tissue in which responses are assessed.

time

Exposure time (h).

biomarker__isoform

Transcript as presented in the publication.

biomarker

Transcript as categorized into one of six main types (*cat*, *gst*, *hsp70*, *hsp90*, *mt* or *sod*).

toxicant

Metal used for exposure.

concentration

Exposure concentration ($\mu\text{g/L}$).

response

The response magnitude, on the linear scale. Unless presented on a log-scale, the response data were extracted and are presented on the same scale as presented in the original publication. Responses on a log-scale were converted to the linear scale upon extraction.

se

The response standard error.

sd

The response standard deviation.

control_response

The response magnitude of the corresponding control group, on the linear scale. Unless presented on a log-scale, the response data were extracted and are presented on the same scale as presented in the original publication. Responses on a log-scale were converted to the linear scale upon extraction.

control_se

The response standard error of the corresponding control group.

control_sd

The response standard deviation of the corresponding control group.

n

Sample size of the exposure group.

n_control

Sample size of the corresponding control group.

group_comparisons

Number of exposure groups compared to the same control group.

n_control_adjusted

Sample size of the corresponding control group, corrected for multiple comparisons ($n_control / group_comparisons$).

LRR

Effect size; log response ratio. Calculated according to Eq. 2 in the article.

vLRR

Response variance. Calculated according to Eq. 3 in the article.

Toxic units

The ‘toxic_units.Rmd’ script contains the code used to obtain toxicity data from the U.S. Environmental Protection Agency ECOTOX database. That data are in turn used to assign a median bivalve LC_{50} to each metal for which data are available. The script saves the median LC_{50} data to the ‘toxic_unit.xlsx’ file.

toxic_unit.xlsx

Contains median bivalve LC_{50} data used in the meta-analyses to determine toxic units. The data are obtained from running the ‘toxic_units.Rmd’ script.

Meta-analysis

The ‘models.Rmd’ script contains the code used to run the meta-analyses, and imports data from the ‘dataset.xlsx’ and ‘toxic_unit.xlsx’ files.

Main models

Nine Bayesian hierarchical random effect models were used to test effects of metal exposure on transcriptional responses, both for metal exposure in general and in meta-regressions using TU (‘log_tu’; toxic unit converted to the log scale) and exposure time (‘log_time’; exposure time converted to the log scale) as independent variables, with and without transcript and tissue as additional moderators.

‘mod1.rds’

Overall response model; intercept model.

‘mod2.rds’

Transcript overall response model; transcript (‘biomarker’) as categorical moderator.

‘mod3.rds’

Tissue overall response model; tissue as categorical moderator.

‘mod4.rds’

Overall concentration-dependent response model; toxic unit (‘log_tu’) as continuous moderator.

‘mod5.rds’

Transcript-specific concentration-dependent response model; toxic unit (‘log_tu’) as continuous moderator and transcript (‘biomarker’) as categorical moderator.

‘mod6.rds’

Tissue-specific concentration-dependent response model; toxic unit (‘log_tu’) as continuous moderator and tissue as categorical moderator.

‘mod7.rds’

Overall time-dependent response model; exposure time (‘log_time’) as continuous moderator.

‘mod8.rds’

Transcript-specific time-dependent response model; exposure time (‘log_time’) as continuous moderator and transcript (‘biomarker’) as categorical moderator.

‘mod9.rds’

Tissue-specific time-dependent response model; exposure time (‘log_time’) as continuous moderator and tissue as categorical moderator.

Heterogeneity

I^2 was estimated for each model as a measure of heterogeneity. The script saves estimated I^2 to the ‘heterogeneity.xlsx’ file.

heterogeneity.xlsx

Contains I^2 used to estimate heterogeneity of all models. The data are obtained from running the ‘models.Rmd’ script.

Prediction intervals

Prediction intervals were determined for each effect size in the no moderator/ categorical moderator models (models 1-3). The script saves the lower and upper bound of 95 % prediction intervals to the ‘prediction_intervals.xlsx’ file.

prediction_intervals.xlsx

Contains lower and upper bounds of 95% prediction intervals for no moderator/ categorical moderator models (models 1-3). The data are obtained from running the ‘models.Rmd’ script.

Sensitivity analyses

Correlation factor

For the main models, a variance-covariance matrix was created assuming a correlation factor of 0.5 for multiple responses measured in the same exposure group (‘study’ \times ‘toxicant’ \times ‘concentration’). In the sensitivity analysis, alternative models were run instead assuming correlation factors of 0.1 (‘mod1_01.rds’, ‘mod2_01.rds’, ..., ‘mod9_01.rds’) and 0.9 (‘mod1_09.rds’, ‘mod2_09.rds’, ..., ‘mod9_09.rds’), respectively.

Publication bias

To test the influence of publication bias on the outcome of the meta-analyses, meta-regressions were run based on square roots of inverted effective sample sizes ('sqrt_inv_eff_n'; 'bias_mod1a.rds', 'bias_mod2a.rds', 'bias_mod3a.rds') and on inverted effective sample sizes ('inv_eff_n'; 'bias_mod1b.rds', 'bias_mod2b.rds', 'bias_mod3b.rds'). Also, the relationship between inverted effect size and toxic unit and exposure time, respectively were modeled to examine potential bias ('bias_text_tu.rds' and 'bias_test_time.rds', respectively).

Figures

The 'plots.Rmd' script contains the code used to produce all figures presented in the article and supplementary material. It imports data from the 'dataset.xlsx', 'heterogeneity.xlsx', 'prediction_intervals.xlsx' and 'toxic_unit.xlsx' files as well as R objects (models) created in the 'models.Rmd' script.