# Methods

### *Effect size*

To quantify the impacts of developmental stress on mitochondrial function we calculated the standardized mean difference, assuming heteroscedastic population variances () (Bonett, 2008, 2009) along with its associated sampling error. We used the *escalc* function of the *metafor* package (vers. 4.8.0) (Viechtbauer, 2010) in R (vers. 4.5.1) which implements a correction to account for possible bias resulting from effect sizes calculated with small sample sizes (Borenstein & Hedges, 2009). We choose to use as our effect size because our data: 1) contained substantial amounts of ratio data (e.g., respiratory control ratios, RCR, relative gene expression) which makes interpretation with alternative effect size measures, such as log response ratio, challenging; 2) percentages and zero measurement variables and 3) skewed measurement variables. is more robust to these types of measurement variables.

Given the diversity of measurement variables we paid particular attention to effect size direction to ensure that positive increase of means all represent increased in mitochondrial function. ONDI ELABORATES AND TABLE

We calculated as the mean difference between control and treatment groups divided by the pooled standard deviation. As such, positive effect sizes represent situations where the mean of the experimental group was larger than the control group. Again, we corrected the direction of effect size (multiplied by -1) to ensure their interpretation remained consistent with respect to mitochondrial function.

### *Meta-Analysis*

We analysed our data using multilevel meta-analytic (MLMA) and meta-regression (MLMR) models in R using *metafor* (vers. 4.8.0 Viechtbauer, 2010). Meta-analytic mean estimates and meta-regression models were plotted using orchard plots which convey overall meta-analytic means, alongside raw effect sizes and prediction intervals, which describe effect heterogeneity (vers. 2.0, Nakagawa *et al.*, 2021, 2023).

#### Multi-level Meta-analysis (MLMA) Models

Using data subsets for each of our major stressor categories, we first fit multi-level meta-analysis (MLMA) models (i.e., intercept-only models) with , that included study, tissue type, and phylogeny as random effects to account for non-independence and identify sources of variability. We also explored models that included species and sample dependency as random effects but these were highly confounded with phylogeny and study so we simplified models by removing these terms. Given *metafor* does not estimate a residual variance by default we also included an observation-level random effect in our models. The size of datasets varied in the number of species and tissues ([1](#tbl-heterogeneity)). As such, we simplified the random effect structure for some datasets; including only random effects with six or more levels. Our MLMA models allowed us to partition variation in among these key sources while accounting for total sampling variance. We took a pluralistic approach to heterogeneity reporting, by reporting absolute (), relative (i.e., , *sensu* Nakagawa & Santos (2012); Noble *et al.* (2022)) and magnitude measures (both and ) of heterogeneity as suggested by Yang *et al.* (2023). We also report each of their stratified versions for each random effect level (Nakagawa & Santos, 2012; Yang *et al.*, 2023). We also present 95% prediction intervals which describe the expected distribution of effects from future studies (Nakagawa *et al.*, 2021; Noble *et al.*, 2022).

A phylogeny was derived by first cross-checking taxa names using the Open Tree of Life (OTL) with the *rotl* package in R (vers. 3.1.0) (Michonneau, Brown, & Winter, 2016), and plotting taxa using *ggtree* to visualize the tree (vers. 3.16.0) (Yu *et al.*, 2017). We resolved any taxa names that were outdated or changed. Once names in the dataset and *rotl* database matched we exported species names and built a time calibrated phylogeny for the species in our dataset using TimeTree [timetree.org; Kumar *et al.* (2022)]. We used the R packages *ape* (vers. 5.8.1) (Paradis & Schliep, 2019) and *phytools* (vers. 2.4.4) (Revell, 2012) to prune the tree for individual analyses and calculate phylogenetic correlation matrices used in meta-analytic models.

#### Multi-level Meta-Regression (MLMR) Models

After quantifying levels of heterogeneity in each of the different stressor datasets, we fit a series of multi-level meta-regression (MLMR) models to test our key questions. In all models, we included the same random effects as we used in our MLMA models, however, MLMR models differed in the moderators included because each stressor type had unique sources of heterogeneity that we expected would explain effect variance. Despite this, a number of moderators were relevant to all stressor manipulation types and were included in all models. These included: 1) measurement trait category (i.e., antioxidants, oxidative damage etc.), 2) stage of manipulation (pre-natal/post-natal or both) and 3) taxa (i.e., class: mammals, birds, amphibians, fish). Some stressor categories were made up of a single class (e.g., social deprivation was only done in mammals) and so we did not include the taxa moderator in these models. We fit models assuming heteroscedastic variation among levels of a given moderators to capture differences in effect variation.

We had a substantially larger dataset of studies measuring nutritional stress. These studies were more variable and so we included some additional moderators to capture variation in effects. These included: 1) whether the nutritional manipulation involved over or under nutrition and 2) the type of nutritional manipulation done (e.g., high fat, low protein, etc.). We expected that the different measurement categories may exhibit different responses to nutritional stress and so we also included an interaction between measurement category and over/under nutrition. However, we did not have enough data to estimate an interaction between manipulation type and functional category. We included models with nutrition type, under/over nutrition and a combined model with both main effects and evaluated support among models using AIC, choosing the simplest model (fewest estimated parameters) when the AIC values were within 2 units of each other.

# Results

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1- Relative heterogeneity estimates (), sample sizes (), number of studies () and number of species () for each of the developmental stressor categories. Relative heterogeneity is reported as (in percent) based on the intercept-only multi-level meta-analytic model. Total heterogeneity, along with stratified version of (study, phylogeny, tissue type and observation) are also provided. Note we also present other heterogeneity measures in the supplementary material.   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Corticosterone | 140 | 8 | 10 | 86 | 12.2 | 0 | 68 | 6.6 | | Disturbance | 190 | 8 | 16 | 80 | 57.2 | 11 | 0 | 11.8 | | Nutrition | 645 | 15 | 51 | 80 | 8.6 | 13 | 10 | 47.9 | | Social Deprevation | 172 | 2 | 10 | 82 | 36.8 | – | – | 45.1 | |
| **Figure** 1- **Phylogeny of species included in the meta-analysis**. Phylogenetic relationships and divergence times are based on TimeTree [timetree.org; Kumar *et al.* (2022)]. The phylogeny was pruned to only include species in the dataset. |

In total, we collected 1147 effect sizes from 21 species ([Figure 1](#fig-phylo)) and 86 studies across the different stressors (Table [1](#tbl-heterogeneity)). There were clear taxonomic biases. The majority of effects came from studies on mammals (effects = 951, studies = 64), and these were only from 4 species (*Mus musculus*, *Rattus norvegicus*, *Ovis aries* and *Capra hircus*). Fish had the second most effects (105, species = 9 and studies = 10) followed by birds (75, species = 6 and studies = 11) and amphibians (16, species = 2 from only 1 study).

|  |
| --- |
| **Figure** 2- **Overall meta-analytic mean estimates for each developmental stressor category (a) corticosterone, (b) social deprevation, (c) disturbance stress and (d) nutritional stress**. Mean estimates are shown as large points, with 95% confidence (thick black bars) and prediction (thin black bars). The number of effects (k) along with the number of studies (in parantheses) are provided for each stressor category. The dashed line represents no effect (i.e., = 0). Raw data are also shown and are weighted by their inverse sampling error (precision). |

## Corticosterone developmental impacts on mitochondrial function

Overall, exposure to corticosterone during development negatively impacted mitochondrial function ( = -0.92, 95% CI = -1.69 to -0.16, = 0.02; [Figure 2](#fig-orchard_int)a). Effect heterogeneity was high (95% Prediction Intervals: -3.07 to 1.23; Table [1](#tbl-heterogeneity) and see Table S [2](#tbl-heterogeneity_cv) & Table S [3](#tbl-heterogeneity_m2)), with 12.24% of variation being driven by differences among studies (; Table [1](#tbl-heterogeneity)) and 67.53% of variation being driven by differences among tissues (; Table [1](#tbl-heterogeneity)) with little to no variation explained by phylogeny (Table [1](#tbl-heterogeneity)).

Prenatal exposure to corticosterone had a larger negative impact on mitochondrial function than postnatal exposure ( = -1.09, 95% CI = -2.43 to 0.26), however, this effect was not significant ( = 0.11; [Figure 5](#fig-orchard_prepost)a).

Mitochondrial function was supressed across all major functional categories in a similar fashion with no significant differences among functional categories ( = 1.49, = 0.22 [Figure 3](#fig-orchard_measure)a). Treatment with corticostrone supressed respiration and metabolic capacity the greatest (respiraton: = -0.81, 95% CI: -1.58 to -0.05, = 0.04; metabolic capacity: = -1.09, 95% CI: -1.86 to -0.33, = 0.01) and increasing oxidative damage ( = -0.71, 95% CI: -1.49 to 0.07, = 0.07).

While effects appeared to be stronger in mammals, there were no significant differences among classes ( = 0.26, = 0.77; [Figure 4](#fig-orchard_taxa)a). Overall, mitochondrial function was supressed in mammals ( = -1.15, 95% CI: -2.17 to -0.14, = 0.03) and fish ( = -1.15, 95% CI: -2.55 to 0.26, = 0.11), with effects being weaker (albeit still negative) in birds ( = -0.49, 95% CI: -2.01 to 1.02, = 0.52).

|  |
| --- |
| **Figure** 3- **Overall meta-analytic mean estimates across each functional trait category for each developmental stressor type (a) corticosterone, (b) social deprevation, (c) disturbance stress and (d) nutritional stress**. Mean estimates are shown as large points, with 95% confidence (thick black bars) and prediction (thin black bars). The number of effects (k) along with the number of studies (in parantheses) are shown provided for each trait category. The dashed line represents no effect (i.e., = 0). Raw data are also shown and are weighted by their inverse sampling error (precision). |

## Social deprivation developmental impacts on mitochondrial function

Social deprivation during development had a negative impact on mitochondrial function ( = -0.8, 95% CI = -1.4 to -0.21, = 0.01; [Figure 2](#fig-orchard_int)b). Effect heterogeneity was again high (95% Prediction Intervals: -3.43 to 1.82; Table [1](#tbl-heterogeneity); and see Table S [2](#tbl-heterogeneity_cv) & Table S [3](#tbl-heterogeneity_m2)), with 36.82% of the variation being driven by differences among studies () (Table [1](#tbl-heterogeneity)).

Social deprevation studies were only ever done with mammals, and all studies manipulated environments postnatally. Mitochondrial function was supressed across all major functional categories but they did not differ significantly from each other ( = 1.7, = 0.19; [Figure 3](#fig-orchard_measure)b). Effects were more negative for antioxidant capacity ( = -1.51, 95% CI: -2.48 to -0.53, = 0) and oxidative damage ( = -1.17, 95% CI: -2.03 to -0.32, = 0.01). While metabolic capacity was also supressed overall, it was not significantly different from zero ( = -0.47, 95% CI: -1.03 to 0.1, = 0.11; [Figure 3](#fig-orchard_measure)b).

|  |
| --- |
| **Figure** 4- **Overall meta-analytic mean estimates for each taxonomic group (class) across developmental stressors (a) corticosterone, (b) social deprevation, (c) disturbance stress and (d) nutritional stress**. Mean estimates are shown as large points, with 95% confidence (thick black bars) and prediction (thin black bars). The number of effects (k) along with the number of studies (in parantheses) are shown provided for each class. The dashed line represents no effect (i.e., = 0). Raw data are also shown and are weighted by their inverse sampling error (precision) |

## Disturbance developmental impacts on mitochondrial function

Disturbance during development had a negative impact on mitochondrial function ( = -0.8, 95% CI = -1.54 to -0.06, = 0.03; [Figure 2](#fig-orchard_int)c). Effect heterogeneity was high (95% Prediction Intervals: -3.17 to 1.57; Table [1](#tbl-heterogeneity)), with substantial variation across studies ( = 57.25%)(Table [1](#tbl-heterogeneity)).

Impacts of disturbance affected functional categories significantly differently ( = 3.18, = 0.04 [Figure 3](#fig-orchard_measure)c). Oxidative damage was most impacted by disturbance stress ( = -1.02, 95% CI: -1.92 to -0.12, = 0.03), and differed significantly from antioxidant ( = -0.62, 95% CI: -1.49 to 0.25, = 0.16) and metabolic capacity ( = -0.53, 95% CI: -1.42 to 0.36, = 0.24) ([Figure 3](#fig-orchard_measure)c).

Disturbance stress occured only postnatally in fish and birds, whereas in mammals disturbance stressors were also applied prenatally. As such, there was some taxonomic bias. However, we did not find any significant differences among classes ( = 1.48, = 0.23; [Figure 4](#fig-orchard_taxa)b). Despite this, we found evidence that prenatal and post-natal disturbance stressors had different impacts on mitochondrial function. Prenatal disturbance stressors had a larger negative (albeit non-significant) impact on mitochondrial function compared to postnatal disturbance (Contrast = -0.45, 95% CI = -1.55 to 0.65, = 0.42; [Figure 5](#fig-orchard_prepost)b). Overall, prenatal disturbance stressors had a larger negative impact on mitochondrial function ( = -1.19, 95% CI = -2.07 to -0.31, = 0.01), whereas postnatal disturbance stressors had a smaller negative impact ( = -0.74, 95% CI = -1.46 to -0.02, = 0.04) ([Figure 5](#fig-orchard_prepost)b).

|  |
| --- |
| **Figure** 5- **Overall meta-analytic mean estimates different stages of manipulation (pre-natal, post-natal or both) across different developmental stressors (a) corticosterone, (b) social deprevation, (c) disturbance stress and (d) nutritional stress**. Mean estimates are shown as large points, with 95% confidence (thick black bars) and prediction (thin black bars). The number of effects (k) along with the number of studies (in parantheses) are shown provided for each stage. The dashed line represents no effect (i.e., = 0). Raw data are also shown and are weighted by their inverse sampling error (precision) |

## Nutrition developmental impacts on mitochondrial function

Nutritional stress during development had a weaker negative impact on mitochondrial function overall ( = -0.38, 95% CI = -1.03 to 0.27, = 0.25; [Figure 2](#fig-orchard_int)d). Effect heterogeneity was high (95% Prediction Intervals: -2.85 to 2.08; Table [1](#tbl-heterogeneity); and see Table S [2](#tbl-heterogeneity_cv) & Table S [3](#tbl-heterogeneity_m2)), with variation being driven by differences between studies ( = 8.57%), phylogeny ( = 13.21%) and tissue ( = 10.29%) (Table [1](#tbl-heterogeneity)).

Importantly, there was a taxonomic bias in the types of nutritional manipulations done, with fish studies being the only ones to manipulate carbohydrates; only total food was manipulated across all classes. Under- or over-nutrition did not significantly differ in their impacts on mitochondrial function ( = 0.02, = 0.9; [Figure 6](#fig-orchard_nutrition)a) and there was weak evidence for an interaction between measurement type and under/over-nutrition ( = 1.4, = 0.2). In fact, there was little evidence that mitochodnrial function differed across functional trait categories generally ([Figure 3](#fig-orchard_measure)d). The best supported model, suggested, protein and fat manipulations did have a larger negative impact on mitochondrial function compared to carbohydrate manipulations but this effect was confounded with taxonomic class where only mammals had protein manipulations (Protein Contrast = -1.32, 95% CI: -2.54 to -0.1, = 0.03; Fat Contrast = -1.1, 95% CI: -2.31 to 0.1, = 0.07 [Figure 6](#fig-orchard_nutrition)b). There was no evidence that the stage of manipulation when nutritional stress took place impacted mitochondrial function ([Figure 5](#fig-orchard_prepost)d).

|  |
| --- |
| **Figure** 6- **Overall meta-analytic mean estimates for nutritional stressors (a) under- and over-nutrition, (b) nutrition type and (c) nutrition type and over/under-nutrition**. Mean estimates are shown as large points, with 95% confidence (thick black bars) and prediction (thin black bars). The number of effects (k) along with the number of studies (in parantheses) are shown provided for each stressor type. The dashed line represents no effect (i.e., = 0). Raw data are also shown and are weighted by their inverse sampling error (precision) |

# References

Bonett, D.G. (2008) Confidence intervals for standardized linear contrasts of means. *Psychological Methods* **13**, 99. American Psychological Association.

Bonett, D.G. (2009) Meta-analytic interval estimation for standardized and unstandardized mean differences. *Psychological methods* **14**, 225. American Psychological Association.

Borenstein, M. & Hedges, L.V. (2009) Effect sizes for meta-analysis. *In: The Handbook of Research Synthesis and Meta-analysis (eds Cooper H, Hedges LV, Valentine JC). Russell Sage Foundation, New York.*, 207–243.

Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G. & Hedges, S.B. (2022) TimeTree 5: An expanded resource for species divergence times. *Molecular biology and evolution* **39**, msac174. Oxford University Press.

Michonneau, F., Brown, J.W. & Winter, D.J. (2016) Rotl: An R package to interact with the open tree of life data. *Methods Ecol. Evol.* **7**, 1476-1481. doi:10.1111/2041-210X.12593.

Nakagawa, S., Lagisz, M., O’Dea, R.E., Pottier, P., Rutkowska, J., Senior, A.M., Yang, Y. & Noble, D.W. (2023) orchaRd 2.0: An r package for visualising meta-analyses with orchard plots. *Methods in Ecology and Evolution* **14**, 2003–2010. Wiley Online Library.

Nakagawa, S., Lagisz, M., O’Dea, R.E., Rutkowska, J., Yang, Y., Noble, D.W. & Senior, A.M. (2021) The orchard plot: Cultivating forest plots for use in ecology, evolution and beyond. *Research Synthesis Methods* **12**, 4–12.

Nakagawa, S. & Santos, E.S.A. (2012) Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* **26**, 1253–1274. Springer.

Noble, D.W., Pottier, P., Lagisz, M., Burke, S., Drobniak, S.M., O’Dea, R.E. & Nakagawa, S. (2022) Meta-analytic approaches and effect sizes to account for ‘nuisance heterogeneity’ in comparative physiology. *Journal of Experimental Biology* **225**, jeb243225.

Paradis, E. & Schliep, K. (2019) Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528.

Revell, L.J. (2012) Phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223.

Viechtbauer, W. (2010) Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**, 1–48. URL: https://www.jstatsoft.org/v36/i03/.

Yang, Y., Noble, D.W., Spake, R., Senior, A.M., Lagisz, M. & Nakagawa, S. (2023) Measuring biological generality in meta-analysis: A pluralistic approach to heterogeneity quantification and stratification. *EcoEvoRxiv*.

Yu, G., Smith, D., Zhu, H., Guan, Y. & Lam, T.T.-Y. (2017) Ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36, doi:10.1111/2041–210X.12628.

# Supplementary Materials

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2- Heterogeneity estimates using CV for developmental stressor datasets   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Corticosterone | 140 | 8 | 10 | 1.1 | 0.41 | 0.00 | 0.97 | 0.30 | | Disturbance | 190 | 8 | 16 | 1.4 | 1.21 | 0.53 | 0.00 | 0.55 | | Nutrition | 645 | 15 | 51 | 3.2 | 1.03 | 1.28 | 1.13 | 2.44 | | Social Deprevation | 172 | 2 | 10 | 1.6 | 1.08 | – | – | 1.19 | |
| Table 3- Heterogeneity M2   |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | Corticosterone | 140 | 8 | 10 | 0.52 | 0.20 | 0.00 | 0.46 | 0.14 | | Disturbance | 190 | 8 | 16 | 0.59 | 0.50 | 0.22 | 0.00 | 0.23 | | Nutrition | 645 | 15 | 51 | 0.76 | 0.25 | 0.31 | 0.27 | 0.59 | | Social Deprevation | 172 | 2 | 10 | 0.62 | 0.41 | – | – | 0.46 | |