Methodology

**1.1 Data Collection**

Relevant papers were identified and recorded using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher et al. 2015). Database searches were conducted between the 9th of October and the 20th of November 2020 on the Web of Science, PubMed, and the University of Glasgow Library using the keywords, “Oxidative Stress” and “Amphibians”. Articles published between 2015-2020 were reviewed and as shown in the PRISMA Flow Diagram (Figure 1), 553 studies were generated through the database search and 3 were identified through searching the reference sections of the papers reviewed. After duplicates were removed, the title and abstract were screened and 85 articles were identified as potentially suitable for the meta-analysis. From the studies obtained, they could only be included in the study if they met the following criteria:

1. The study reported the means, variability (standard deviation (SD) or standard error (SE)) and the sample sizes for the control group (group not exposed to pollutant) and the treatment group (group exposed to pollutant)
2. The study used an appropriate biomarker including the enzymatic biomarkers (SOD, GPx, CAT, GR), non-enzymatic biomarker (GSH) and indicator of oxidative damage (MDA, TBARS)
3. The study reported the amphibian’s development stage (Embryo, Tadpole, Adult)
4. The duration of the experiment was reported
5. The study did not test the effect of a pollutant in combination with another factor.

A total of 30 studies met the criteria for inclusion and from these, a dataset was compiled consisting of 522 estimates. From each study, the mean, SD and sample size were derived from both control and treatment groups, and results presented as SE were transformed into SD using the formula:

SD = SE\*

Most of the data was presented graphically, therefore, numerical data was obtained using digitalising software, WebPlotDigitiser Version 4.4, and this tool has been shown to be a valid and reliable method of data extraction for meta-analyses (Drevon et al. 2017; Rohatgi 2020). From each study, additional data was collected including the type of species, development stage, location of the study, experimental venue (lab or field), pollutant class (Herbicide, Pesticide, Wastewater contaminant, Fungicide, Insecticide, Heavy Metal), biological matrix and the specific biomarker measured. To account for the variation in the pollutant concentrations across the 30 studies, the expected environmental concentration (EEC) ranges were extracted from relevant research papers and inserted into a table (see Appendix 1). Pollutant concentrations were classed into 'Low', 'Medium', and 'High' and Low was the EEC, Medium was slightly above the EEC, and High included concentrations that are not ecologically relevant.

Diagram

Description automatically generated

**Figure 1: PRISMA Flow Diagram outlining literature review process.**

**1.2 Meta-Analysis**

The data collected was entered into a excel spreadsheet and then imported to Version 4.0.4 of R and Version 1.4.1106 of R studio and the ‘*metafor’, ‘ggplot2’* and ‘*orchaRd’* packages were used for the meta-analysis (Nakagawa et al. 2020; Viechtbauer 2010; Villanueva & Chen 2019). The ‘*metafor’* package calculated effect sizes and sampling variances using the ‘*escalc’* function and a random effect size model was generated using the ‘*rma.mv’* function. The *‘ggplot2’* and ‘*orchaRd’* packages were used to visualise the results in a forest-like/orchard plot, which shows the overall mean effects, 95% confidence intervals, CIs, and 95% prediction intervals, PIs (Nakagawa et al. 2020). Meta-analyses regularly use forest plots, which display individual point estimates, CIs and an overall mean of effect sizes at the bottom of the graph. However, this study has >500 effect sizes, therefore, the orchard plot was used to clearly visualise the results by displaying the average effect size. Also, unlike forest plots, orchard plots include PIs, which incorporate heterogeneity and predict the range in which the effect size is expected to fall in future studies and effect sizes scaled by their precision are displayed in the graph.

**1.3 Calculating effect sizes**

Effect sizes are a standardised measure of the magnitude of observed effect and, therefore, studies that have measured different pollutants or used different scales of measurement can be directly compared. The means, SDs and sample sizes were used to calculate standard mean differences (SMD), also known as Hedges’ g, using the ‘escalc’ function in the ‘*metafor’* package (Hedges & Vevea 1998). Hedges’ g corrects for possible bias in studies with a small sample size (<20) and as described by Cohen (1988), an effect size of 0.2, 0.5 and 0.8 can be classed as ‘small’, ‘medium’ and ‘large’, respectively. Most studies reported more than one biomarker of oxidative stress, therefore, a total of 522 effect sizes were computed from the 30 studies.

**1.4 Heterogeneity and random-effects model**

The publications reviewed in this study are prone to sources of heterogeneity, including pollutant concentration levels, the type of pollutant, the development stage, species studies, the biological matrix studied and the biomarkers of oxidative stress. The magnitude of heterogeneity between studies was evaluated using the I2 Index, and I2 values of 25%, 50%, and 75% were considered as low, moderate, and high heterogeneity, respectively (Huedo-Medina et al. 2006). Significant heterogeneity was detected in this study (I2 > 75%), therefore, a random effects model was created using the ‘rma.mv’ function in the ‘*metafor’* package. This model was chosen over a fixed-effects model because the true effect sizes in this study are affected by both within-study and between-study heterogeneity, however, fixed-effects models only consider within-study variation (Viechtbauer 2010). A multivariate linear mixed effects model with restricted maximum likelihood estimation (REML) was used to test whether the fixed factors were significantly different from zero, and in this study the fixed factors were biomarker type (SOD, CAT, GPx, GR, GSH, MDA, TBARS) and development stage (Embryo, Tadpole, Adult). Also, Study ID was included as a random effect in all models conducted during this study. Overall effect sizes were visually assessed using an orchard plot, which shows overall mean effects, CIs and PIs. Where the 95% CIs did not cross over zero, the effect was significant (p<0.05).

**1.5 Publication Bias**

Furthermore, studies with significant results are more likely to be published than those with no significant results, resulting in publication bias. Therefore, the possibility of publication bias across the entire study was visually assessed using a funnel plot and quantitively assessed using an Egger's test. If publication bias was detected by the Egger’s test, the trim and fill method was used to calculate the missing studies and adjust the summary effect size (Duval S. & Tweedie R. 2000; Sterne & Egger 2001). Also, publication bias is not the only factor that can create funnel plot asymmetry, because factors associated with both sample size and the effect size, such as poor study design can cause asymmetry (Peters et al. 2008). Therefore, an additional extension known as the contour enhanced funnel plot was used because it displays areas of statistical significance on the plot for each individual study.

Results

**2.1 Publication Bias and heterogeneity**

A visual assessment of the contoured funnel plot (Figure 1) determined that the data points were symmetrical, however, the studies were located within the light grey areas of statistical significance, which indicates that there is publication bias. Egger’s multivariate regression was significantly different from zero (z= -9.39, p<0.0001) confirming the publication bias of the whole data set. The total heterogeneity within this dataset was statistically significant (I2=98.89%, 95% CI [-0.9008 -0.2999], Q=10472.1730(df = 521)).

**A picture containing text, boat, different

Description automatically generated**

**Figure 1: Contour enhanced funnel plot.** Funnel plots of publication bias for a meta-analysis of the effects of pesticides, fungicides, heavy metals, herbicides and water contaminants on biomarkers of oxidative stress. White and dark grey show areas of statistical non-significance and the two areas of lighter grey show areas of statistical significance.

**2.2 Overall dataset**

There were 21 different species of amphibian included in this study (8 Families) and the tadpole stage of development was most frequently used in the experiments (421/ 522 effect sizes) compared to embryos and adults (24 and 77/522 effect sizes, respectively). Most studies used herbicides as the treatment group (171 effect sizes), compared to heavy metal (135 effect sizes) and pesticide (117 effect sizes) exposure. A lower number of studies exposed amphibians to water contaminants (51 effect sizes) and fungicides (54 effect sizes). Enzymatic biomarkers accounted for 353 effect sizes compared to non-enzymatic and indicator biomarkers with 82 and 93 effect sizes, respectively. In this study “Low” and “High” pollutant concentrations each accounted for 232 effect sizes and the “Medium” pollutant concentration included 64 effect sizes.

**2.3 Overall effects of pollution on biomarkers of oxidative stress**

The overall effect of pollutant exposure on biomarkers of oxidative stress was negative, therefore it favoured the treatment group (Figure 2A). The effects of Biomarkers GPx (Hg = -1.05, 95% CI [-1.97 -0.12], p<0.05), MDA (Hg = -0.97, 95% CI [-1.82-0.12], p<0.05), and SOD (Hg = -0.65, 95% CI [-1.28-0.027], p<0.05) were Medium to High and significantly different from zero. Additionally, the effect of pollution on CAT, GR, GSH and TBARS was Small to High and since the CI’s overlapped zero and, because p>0.05, they were not significantly different.

**2.4 Pesticides-Biomarker effects**

As demonstrated in Figure 2B, the exposure of pesticides on amphibians had a significant impact on GPx (Hg = -0.63, 95% CI [0.74-4.15], p<0.05) and MDA (Hg = -1.83, 95% CI [-2.93-0.72], p<0.05). The mean effect size of GPx is negative, so it favours the treatment group. Conversely, the effect size of MDA is positive, which favours the control group. Exposure to pesticides did not have a significant effect on CAT, GSH, TBARS and SOD. In comparison to the other biomarkers, TBARS is poorly represented in this sample because there is only a single effect size.

**2.5 Fungicide-Biomarker effects**

Fungicide exposure (Figure 2C) had a significant impact on SOD (Hg = -8.20, 95% CI [-10.83-5.57], p<0.05), GSH (Hg = -5.77, 95% CI [-10.32-1.23], p<0.05), GPx (Hg =-11.55, 95% CI [-16.11-7], p<0.05), and CAT (Hg = -3.85, 95% CI [-6.95-0.75], p<0.05) and each of these negative effect sizes favoured the treatment group. MDA was not significantly affected by fungicides in this study (p>0.05). Furthermore, exposure to heavy metals (Figure 2D) significantly affected CAT (Hg =-1.89, 95% CI [-3.03-0.75], p<0.05) and GSH (Hg =1.47, 95% CI [0.09-2.84], p<0.05) and both biomarkers favoured the treatment. The biomarkers GR, SOD and GPx were not significantly impacted by exposure to heavy metals (p>0.05).

**2.6 Herbicide-Biomarker effects**

Exposure to herbicides (Figure 2E) significantly impacted GR (Hg =-2.05, 95% CI [-3.27-0.82], p<0.05), CAT (Hg =0.49, 95% CI [0.17-0.81], p<0.05) and GPx (Hg =1.82, 95% CI [0.84-2.79], p<0.05). GR favoured the treatment, and the mean effect sizes of CAT and GPx are positive and favour the control. Exposure to the herbicides did not significantly impact GSH, MDA, SOD and TBARS (p>0.05).

**2.7 Wastewater contaminants- Biomarker effects**

Furthermore, GSH was significantly impacted by Wastewater contaminants (Hg =-2.62, 95% CI [-4.93-0.31], p<0.05) and favoured the treatment (Figure 2F). On the other hand, biomarkers, CAT, GR, TBARS, SOD and MDA were not significantly impacted by the wastewater contaminants. Also, although they were not significantly impacted, CAT (Hg = 1.94, 95% CI [-0.02-3.91]) and SOD (Hg = 0.84, 95% CI [-1.22-2.90]), are both positive and favour the control group.

1. Overall effect
2. Pesticide

Chart, scatter chart

Description automatically generatedChart

Description automatically generated with medium confidence

1. Fungicide
2. Heavy metals

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

1. Wastewater Contaminants
2. Herbicide

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

**Figure 2: Orchard plot of the effect of different pollutant types on the oxidative stress biomarkers.** These orchard plots show the mean effect size (Standard mean difference, SMD), CI (bold line), PI (thin line) and the individual effect sizes and their precision for oxidative stress biomarkers. Where the 95% CIs did not cross over zero, the effect was significant (p<0.05). The horizontal dashed line indicates no effect and negative effect sizes favour the treatment. A illustrates the overall effect of all pollutants on the specific biomarkers of oxidative stress. B, C, D, E, and F visualise the effect sizes of specific biomarkers of oxidative stress exposed to pesticides, fungicides, heavy metals, herbicides and wastewater contaminants, respectively.

**2.8 Overall effects of pollution on the developmental stage**

The overall effect of pollutant exposure on the developmental stage was positive for embryos and adults, favouring the control group, and negative for tadpoles which favoured the treatment group (Figure 2A). The effects of Developmental stages Embryo (Hg = 0.34, 95% CI [-0.27-1.26], p>0.05) and adults (Hg = 0.49, 95% CI [-0.27-1.26], p>0.05) were Low to High and non-significantly different from zero since the CI’s overlapped zero. The effect of pollution on Tadpoles (Hg = -0.87, 95% CI [-1.20-(-0.53)], p<0.05) was High and significantly significant. In this analysis, Fungicide is not represented because there was only one developmental stage, Tadpole. Therefore, it has not been analysed separately but is included in this model and Figure 3A.

**2.4 Pesticides-Developmental stage effects**

As shown in Figure 3B, pesticide exposure has a significantly high effect on Tadpoles (Hg = -1.11, 95% CI [-1.73-0.49], p<0.05) and, as it was negative, it favoured the treatment. Pesticide exposure also has a high, significant effect on the Adult developmental stage (Hg = -1.41, 95% CI [0.59-2.23], p<0.05), although it is positive and favours the control. Embryos had a high effect; however, it was not significant (p>0.05).

**2.5 Heavy metals- Developmental stage effects**

The effect of heavy metal exposure on developmental stages showed an overall medium-high negative response (Figure 3C). Although it was not significant for the embryonic (Hg = -0.19, 95% CI [-2.92-2.52], p>0.05), tadpole (Hg = -0.62, 95% CI [-1.32-0.069], p>0.05) or adult (Hg = -0.06, 95% CI [-1.28-1.16], p>0.05) stages. Therefore, heavy metal exposure did not significantly impact the development stages.

**2.6 Herbicide- Developmental stage effects**

In comparison to the tadpoles (160 effect sizes), embryos (6) and adults (2) are very underrepresented in this sample. Although, as shown in Figure 3D, there is an overall positive affect, and Embryos demonstrate a high, significant effect (Hg = 3.69, 95% CI [3.06-4.32], p<0.05). Tadpoles and Adults on the other hand demonstrate a low, non-significant effect in response to herbicides.

**2.7 Wastewater contaminants- Developmental stage effects**

Similarly to herbicide exposure, embryos (6 effect sizes) and adults (6) are underrepresented in this sample in comparison to tadpoles (37). As demonstrated in Figure 3E, embryos and adults have a negative effect, favouring the treatment and tadpoles have a positive effect, favouring the control. The water contaminants exposure has low to high effects on developmental stages, and the effect on adults is significant (Hg = -3.28, 95% CI [-6.21-(-0.35)], p<0.05) and the effect on embryos and tadpoles is non-significant.

1. Pesticide
2. Overall Effects

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

1. Herbicide
2. Heavy Metals

Chart

Description automatically generated

1. Wastewater Contaminants

**Figure 3: Orchard plot of the effect of different pollutant types on the developmental stage.** These orchard plots show the mean effect size (Standard mean difference, SMD), CI (bold line), PI (thin line) and the individual effect sizes and their precision for developmental stages. Where the 95% CIs did not cross over zero, the effect was significant (p<0.05). The horizontal dashed line indicates no effect and negative effect sizes favours the treatment. A shows the overall effect of all pollutants on the specific developmental stages. B, C, D, and E visualise the effect sizes of specific biomarkers of oxidative stress exposed to pesticides, heavy metals, herbicides and wastewater contaminants, respectively.

**References**

Cohen, J. 1988, Statistical power analysis for the behavioral sciences, 2nd edn, Psychology Press, London;New York

Drevon, D., Fursa, S.R. & Malcolm, A.L. 2017, "Intercoder Reliability and Validity of WebPlotDigitizer in Extracting Graphed Data", Behavior modification, vol. 41, no. 2, pp. 323-339.

Duval S. & Tweedie R. 2000, "Trim and Fill: A Simple Funnel-Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis", Biometrics, vol. 56, no. 2, pp. 455-463.

Egger, M., Smith, G.D., Schneider, M. & Minder, C. 1997, "Bias in meta-analysis detected by a simple, graphical test", BMJ, vol. 315, no. 7109, pp. 629-634.

Hedges, L.V. & Vevea, J.L. 1998, "Fixed- and Random-Effects Models in Meta-Analysis", Psychological methods, vol. 3, no. 4, pp. 486-504.

Huedo-Medina, T.B., Sanchez-Meca, J., Marin-Martinez, F. & Botella, J. 2006, "Assessing heterogeneity in meta-analysis : Q statistic or I2 index?", Psychological methods, vol. 11, no. 2, pp. 193-206.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L.A. & PRISMA-P Group 2015, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement", Systematic reviews, vol. 4, no. 1, pp. 1-1.

Nakagawa, S., Lagisz, M., O'Dea, R.E., Rutkowska, J., Yang, Y., Noble, D.W.A. & Senior, A.M. 2020, "The orchard plot: Cultivating a forest plot for use in ecology, evolution, and beyond", Research synthesis methods, vol. 12, no. 1, pp. 4-12.

Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R. & Rushton, L. 2008, "Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry", Journal of clinical epidemiology, vol. 61, no. 10, pp. 991-996.

Rohatgi, A. 2020, 4th edition. Available from: [https://automeris.io/WebPlotDigitizer](about:blank)

Sterne, J.A.C. & Egger, M. 2001, "Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis", Journal of clinical epidemiology, vol. 54, no. 10, pp. 1046-1055.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of statistical software, vol. 36, no. 3, pp.1-48.

Villanueva, R.A.M. and Chen, Z.J., 2019. ggplot2: Elegant graphics for data analysis.