Running title: persistent and non-negligible decline effects

Decline effects are rare in ecology: Comment

Yefeng Yang[1,2,3,*], Malgorzata Lagisz[1,4], Shinichi Nakagawa[1,4]

1. Evolution & Ecology Research Centre and School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia

2. Department of Biosystems Engineering, Zhejiang University, Hangzhou 310058, China

3. Department of Infectious Diseases and Public Health, Jockey Club College of Veterinary Medicine and Life Sciences, City University of Hong Kong, Hong Kong, China

* Corresponding author. E-mail: yefeng.yang1@unsw.edu.au (YY)

4. These authors supervised this work equally and are joint senior authors

ORCID

Yefeng Yang: 0000-0002-8610-4016

Malgorzata Lagisz: 0000-0002-3993-6127

Shinichi Nakagawa: 0000-0002-7765-5182

(**INTRODUCTION**)

Recently, Costello and Fox (2022) tested, with a large dataset, the hypothesis of whether there is a widespread decline effect in the discipline of ecology. In other words, the magnitude of the reported ecological effect sizes declines over time (Leimu and Koricheva 2004). Contrary to early results from much smaller datasets (Jennions and Møller 2002, Barto and Rillig 2012), Costello and Fox (2022), using 466 ecological meta-analyses with > 100,000 effect sizes, concluded that there was no systematic decline effect across the field of ecology – only ~5% of ecological meta-analyses showed statistical evidence of a decline effect. This conclusion seems to be "good news" and has important field-wide implications. For example, the temporal stability of the cumulative evidence can alleviate the concerns about policy-making for conservation and environmental management (Koricheva and Kulinskaya 2019).

However, we point out that for their test, Costello and Fox (2022) employed a procedure akin to vote-counting, which is not a preferred method for assessing cumulative evidence in any discipline (Freemantle and Geddes 1998, Combs et al. 2011, Harrison 2011, Gurevitch et al. 2018). Therefore, we have re-analysed their large dataset using a second-order meta-analysis or meta-meta-analysis, which uses a meta-analytic model to statistically synthesize results across different meta-analyses (Fanelli et al. 2017, Nakagawa et al. 2019), and came to a different conclusion. In the Comment, we first report our results by comparing them with those of Costello and Fox (2022). Second, we explain the limitations of the vote-counting method in identifying the decline effect. Finally, to facilitate testing of the decline effect, we provide recommendations on how to conduct and report the decline effect test in ecological meta-analyses (all the code can be found in https://github.com/Yefeng0920/decline_effect_Ecology).

DECLINE EFFECTS ARE PERSISTANT AND NOT NEGLIGIBLE IN ECOLOGY

As mentioned by Costello and Fox (2022), many methods are available for the identification of the decline effect in a given meta-analysis. From among these, they chose two formal statistical methods. The first is a correlation-based approach, which they acknowledge has many limitations but is still used in order to be comparable to previous results (Jennions and Møller 2002). In contrast, Costello and Fox (2022) preferred the second approach – the regression-based approach, which consists of two steps. First, they regressed effect size estimates on publication year ($year_i$) for each meta-analysis (i.e., univariate meta-regression with $year_i$ as a predictor; Nakagawa and Santos 2012, Koricheva and Kulinskaya 2019), where p-value of $year_i$'s slope (β_{year}) < 0.05 indicates that the examined meta-analysis shows the statistical evidence of a decline effect. Second, they used a vote-counting-like method to compute the percentage of p-value < 0.05 for β_{year} that were obtained from the first step. Under the binomial distribution, they tested the percentage of the p-values against the null hypothesis and concluded that the decline effects were rare at the nominal level of 0.05 in the field of ecology (details see Costello and Fox 2022).

Rather than using the percentage of statistically significant β_{year} obtained from the first step, our approach treated slope β_{year} as a standardised effect size measure, which has been used in ecology (De Frenne et al. 2013, Morrissey 2016) and other disciplines (see Nieminen et al. 2013 for medical sciences; Peterson and Brown 2005, Rose and Stanley 2005 for social sciences). We fitted a meta-analytic model to quantitively aggregate β_{year} across 466 meta-analyses, weighting estimates by the inverse square of standard error $SE[\beta_{year}]$. Then, we found that there was a statistically significant systematic decline effect in ecology (overall/pooled β_{year} = −0.0034, 95% confidence interval (CI) = −0.0054 to −0.0014; *p*-value = 0.0008; Figure 1; Supplementary Materials).

Importantly, our meta-meta-analysis can produce new insights that the vote-counting method is unlikely to provide. For example, the estimates of β_{year} were consistent across 466 meta-analyses with a small amount of heterogeneity among these slopes ($I^2_{between-MA}$ = 16%; cf. Senior et al. 2016) (Supplementary Materials). This amount of heterogeneity indicates that the decline effect, albeit small, is persistent across ecological meta-analyses. We also found that the types of effect sizes used could explain 9.4% of the variation ($R^2_{marginal}$ = 0.094; Nakagawa and Schielzeth 2013); the ‘global’ decline effect was mainly driven by *Zr* (Fisher’s r-to-z; pooled $\beta_{Zr\sim year}$ = −0.0084, 95% CI = −0.0121 to −0.0047; Figure 1B). Apart from lnRR (log response ratio), standardized mean difference (SMD) and uncommon effect sizes (e.g., odds ratio) tend to decline over time (albeit non-significantly).

Furthermore, Costello and Fox (2022) speculated that the “decline in the decline effect itself” may explain the contrasting results obtained by early researchers (Jennions and Møller 2002), who found a ubiquitous decline effect in ecological meta-analyses published before 2002. We statistically tested this effect by using a univariate meta-regression with the publication year of meta-analysis papers as a predictor. We found that there has been no statistical evidence of the decline in the decline effect (i.e., estimates of decline effect are not related to publication year of meta-analyses) (Figure 1C).

Taken together, we concluded that there is indeed a consistent ‘global’ decline effect in ecology although the magnitude of the decline effect is small. However, this effect is certainly not negligible, with the estimates of *Zr* being exaggerated by an average of 0.13 units (equivalent to Cohen’s “small” effect size, Cohen 1988; Supplementary Materials).

THE ABSENCE OF EVIDENCE OF THE DECLINE EFFECTS IS NOT THE EVIDENCE OF THE ABSENCE

Costello and Fox (2022) acknowledged that the sample size of the most meta-analyses in their dataset was very small (< 5 effect sizes). Therefore, we expected that the reason why they failed to find the systematic decline effect in ecology is that the vote-counting they used is underpowered, namely reducing the likelihood of detecting a decline effect when it exists (committing type II errors). Indeed, we found that only ~3% of the meta-analyses had a statistical power equal or over the nominal power (80%) to detect a decline effect, with a median power of 17% (0.17, 95% CI = 0.16 to 0.19; Figure 2A). In contrast, the power of our meta-meta-analytic approach was as high as 92%. In addition to the issue of statistical power, many researchers have criticized other flaws in the vote-counting method (Harrison 2011, Nakagawa and Poulin 2012, Koricheva and Gurevitch 2014). For example, it (i) mainly focuses on statistical significance, whose drawbacks have been highlighted elsewhere (Nakagawa and Cuthill 2007, Wasserstein and Lazar 2016, Amrhein et al. 2019); (ii) is incapable of estimating the magnitude and precision of the systematic decline effect; (ii) cannot adjust for the impact of decline effect on the estimation of effect sizes (bias-adjusted effect sizes; cf. Kvarven et al. 2020, Nakagawa et al. 2022, Yang et al. 2022). The meta-meta-analytic approaches can deal with all these issues (see below and Supplementary Materials).

Our meta-meta-analytic approach is a second-order meta-analysis, which uses statistical models to aggregate evidence from multiple first-order meta-analyses (Gurevitch et al. 2018, Nakagawa et al. 2019, Oh 2020). The meta-meta-analytic approach can account for (at least in part) residual sampling variance that the first-order meta-analyses cannot eliminate, thus obtaining a more precise estimate than that in the first-order meta-analyses (Schmidt and Oh 2013). Researchers from many other disciplines have already used second-order meta-analysis to detect the 'global' decline effect (Fanelli et al. 2017, Fanshawe et al. 2017, Pietschnig et al. 2019). Here, we point out two important procedures that we suspect Costello and Fox (2022) may have missed. First, before identifying the decline effect for each meta-analysis, we need to scale different effect size measures, such that the slopes (β_{year}) resulting from different meta-analyses could be directly compared across SMD, lnRR, and *Zr* (Schielzeth 2010, Nakagawa et al. 2017a). Note that we have tested that scaling effect size estimates did not alter the significance estimates for the meta-regression model slope (β_{year}) and thus did not change qualitative conclusions (See Supplementary Materials for a comparison of scaling *vs.* non-scaling effects). Second, we need to flip the signs of β_{year} for meta-analyses with negative overall means before conducting a meta-meta-analysis to aggregate them. For a meta-analysis whose mean/pooled effect size (overall effect) is positive, the sign of β_{year} is expected to be negative if there is a decline effect. In contrast, for a meta-analysis having a negative overall mean, the signs of β_{year} is expected to be positive if there is a decline effect (Figure 2B).

Therefore, when using a meta-analytic model to statistically aggregate β_{year} across different meta-analyses, we need to flip the signs of β_{year} whose associated mean effect sizes are negative. Otherwise, the meta-meta-analysis would create artefactual heterogeneity among β_{year} slopes and thus reduce the statistical power to identify the systematic pattern of the decline effect across ecology. We expect more applications of the meta-meta-analyses in ecology, addressing high-order ecological questions and meta-science questions (the science of science; also known as meta-research; Fidler et al. 2017, Nakagawa et al. 2019).

CONCLUSIONS AND RECOMMENDATIONS

The decline effect can lead to undesirable consequences, such as inflated effect size estimates, temporally unstable evidence being used in ecological policy-making, and even undermining the society’s and scientists’ faith in the use of meta-analytic approaches (Koricheva and Kulinskaya 2019). Such a decline effect has been empirically identified in specific and general areas of ecology (Jennions and Møller 2002, Crystal ‐ Ornelas and Lockwood 2020, Van Klink et al. 2020, Clements et al. 2022). Here, with a much larger dataset compiled by Costello and Fox (2022) and a powerful approach (i.e., meta-meta-analysis), we confirmed previous findings: there is a systematic, non-random, and slow, yet non-negligible, decline effect in ecology (i.e., “global” decline effect; Figure 1A). More importantly, such a “global” decline effect is likely consistent across different subfields of ecology (i.e., different meta-analyses). Our finding sets a “strong prior” expectation that the decline effects are common in ecology. Therefore, the decline effect test should be treated as a routine and mandatory part of ecological meta-analyses (Nakagawa et al. 2017b, Koricheva and Kulinskaya 2019). However, the decline effect test has rarely been performed in the current practices of ecological meta-analyses – two independent surveys both showed that only ~5% of ecological meta-analyses performed the decline effect test (Koricheva and Gurevitch 2014, Nakagawa et al. 2022). The main reason for this low-test rate may be due to the decline effect test being often underpowered (of high Type II error rates), as we have shown. Therefore, we finish with recommendations that could mitigate the issues of low statistical power when testing the decline effect in ecological meta-analyses.

A straightforward and effective way to determine if a decline effect exists is to include the publication year ($year_i$) as a moderator variable in a meta-regression (Nakagawa and Santos 2012, Koricheva and Kulinskaya 2019). The model slope β_{year} provides an intuitive metric to identify if the magnitude of effect sizes declines over time. To maximize the statistical power and control for false positive (Type I error rates), we need to account for other forms of publication bias (e.g., small-study effect) as well as heterogeneity by including important moderator variables (i.e., multilevel multi-moderator meta-regression; Rodgers and Pustejovsky 2021, Nakagawa et al. 2022). Such a multilevel multi-moderator meta-regression can provide a more accurate estimate of the decline effect. Existing software is readily available, and examples of using *rma.mv()* function in *metafor* package are provided in the Supplementary Materials (Viechtbauer 2010). Importantly, given the strong expectation of the decline effect yet limited statistical power to detect it, we need to shift our interpretations from the focus of dichotomous categories (*p*-value < 0.05 meaning decline effect *vs. p*-value > 0.05 meaning no decline effect) toward reporting an accurate estimate of the magnitude, precision, and directionality of the temporal changes in effect sizes which are obtained from the multilevel multi-moderator meta-regression. This is aligned with the statistical philosophy in ecological/biological sciences emphasizing the effect size estimates and their ‘biological’ or ‘practical’ significance rather than statistical significance (Nakagawa and Cuthill 2007, Cumming 2014). We believe that our recommendations can assist in the timely detection of the temporal instability of meta-analytic evidence and make effective allocation of research resources and efforts.

ACKNOWLEDGEMENTS

We thank the Faculty of Science and the office of Deputy Vice-chancellor of Research, UNSW, Sydney for the support to YY and SN. YY was funded by the National Natural Science Foundation of China (NO. 32102597). SN and ML were funded by the Australian Research Council Discovery Grant (DP210100812).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

OPEN RESEARCH

Code is available at Github: https://github.com/Yefeng0920/decline_effect_Ecology

LITERATURE CITED

Amrhein, V., S. Greenland, and B. McShane. 2019. Scientists rise up against statistical significance. Nature **567**:305-307.

Barto, E. K., and M. C. Rillig. 2012. Dissemination biases in ecology: effect sizes matter more than quality. Oikos **121**:228-235.

Clements, J. C., J. Sundin, T. D. Clark, and F. Jutfelt. 2022. Meta-analysis reveals an extreme "decline effect" in the impacts of ocean acidification on fish behavior. PLoS biology **20**:e3001511.

Cohen, J. 1988. Statistical power analysis for the behavioral sciences. 2nd edition. Lawrence Erlbaum., Hillsdale, New Jersey.

Combs, J. G., J. Ketchen, David J, T. R. Crook, and P. L. Roth. 2011. Assessing cumulative evidence within 'macro' research: Why meta‐analysis should be preferred over vote counting. Journal of Management Studies **48**:178-197.

Costello, L., and J. W. Fox. 2022. Decline effects are rare in ecology. Ecology:e3680.

Crystal‐Ornelas, R., and J. L. Lockwood. 2020. Cumulative meta‐analysis identifies declining but negative impacts of invasive species on richness after 20 yr. Ecology **101**:e03082.

Cumming, G. 2014. The new statistics: Why and how. Psychological science **25**:7-29.

De Frenne, P., B. J. Graae, F. Rodríguez‐Sánchez, A. Kolb, O. Chabrerie, G. Decocq, H. De Kort, A. De Schrijver, M. Diekmann, and O. Eriksson. 2013. Latitudinal gradients as natural laboratories to infer species' responses to temperature. Journal of Ecology **101**:784-795.

Fanelli, D., R. Costas, and J. P. Ioannidis. 2017. Meta-assessment of bias in science. Proceedings of the National Academy of Sciences **114**:3714-3719.

Fanshawe, T. R., L. F. Shaw, and G. T. Spence. 2017. A large‐scale assessment of temporal trends in meta‐analyses using systematic review reports from the Cochrane Library. Research Synthesis Methods **8**:404-415.

Fidler, F., Y. E. Chee, B. C. Wintle, M. A. Burgman, M. A. McCarthy, and A. Gordon. 2017. Metaresearch for evaluating reproducibility in ecology and evolution. BioScience **67**:282-289.

Freemantle, N., and J. Geddes. 1998. Understanding and interpreting systematic reviews and meta-analyses. Part 2: meta-analyses. Evidence-Based Mental Health **1**:102-104.

Gurevitch, J., J. Koricheva, S. Nakagawa, and G. Stewart. 2018. Meta-analysis and the science of research synthesis. Nature **555**:175-182.

Harrison, F. 2011. Getting started with meta‐analysis. Methods in Ecology and Evolution **2**:1-10.

Jennions, M. D., and A. P. Møller. 2002. Relationships fade with time: a meta-analysis of temporal trends in publication in ecology and evolution. Proceedings of the Royal Society of London. Series B: Biological Sciences **269**:43-48.

Koricheva, J., and J. Gurevitch. 2014. Uses and misuses of meta‐analysis in plant ecology. Journal of Ecology **102**:828-844.

Koricheva, J., and E. Kulinskaya. 2019. Temporal instability of evidence base: a threat to policy making? Trends in ecology & evolution **34**:895-902.

Kvarven, A., E. Strømland, and M. Johannesson. 2020. Comparing meta-analyses and preregistered multiple-laboratory replication projects. Nature Human Behaviour **4**:423-434.

Leimu, R., and J. Koricheva. 2004. Cumulative meta–analysis: a new tool for detection of temporal trends and publication bias in ecology. Proceedings of the Royal Society of London. Series B: Biological Sciences **271**:1961-1966.

Morrissey, M. B. 2016. Meta‐analysis of magnitudes, differences and variation in evolutionary parameters. Journal of Evolutionary Biology **29**:1882-1904.

Nakagawa, S., and I. C. Cuthill. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biological reviews **82**:591-605.

Nakagawa, S., P. C. Johnson, and H. Schielzeth. 2017a. The coefficient of determination R 2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of the Royal Society Interface **14**:20170213.

Nakagawa, S., M. Lagisz, M. D. Jennions, J. Koricheva, D. Noble, T. H. Parker, A. Sánchez-Tójar, Y. Yang, and R. E. O'Dea. 2022. Methods for testing publication bias in ecological and evolutionary meta-analyses. Methods in Ecology and Evolution **13**:4-21.

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

Nakagawa, S., D. W. Noble, A. M. Senior, and M. Lagisz. 2017b. Meta-evaluation of meta-analysis: ten appraisal questions for biologists. BMC biology **15**:1-14.

Nakagawa, S., and R. Poulin. 2012. Meta-analytic insights into evolutionary ecology: an introduction and synthesis. Evolutionary Ecology **26**:1085-1099.

Nakagawa, S., G. Samarasinghe, N. R. Haddaway, M. J. Westgate, R. E. O'Dea, D. W. Noble, and M. Lagisz. 2019. Research weaving: visualizing the future of research synthesis. Trends in ecology & evolution **34**:224-238.

Nakagawa, S., and E. S. Santos. 2012. Methodological issues and advances in biological meta-analysis. Evolutionary Ecology **26**:1253-1274.

Nakagawa, S., and H. Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed‐effects models. Methods in Ecology and Evolution **4**:133-142.

Nieminen, P., H. Lehtiniemi, K. Vähäkangas, A. Huusko, and A. Rautio. 2013. Standardised regression coefficient as an effect size index in summarising findings in epidemiological studies. Epidemiology, Biostatistics and Public Health **10**:e8854.

Oh, I.-S. 2020. Beyond meta-analysis: Secondary uses of meta-analytic data. Annual Review of Organizational Psychology and Organizational Behavior **7**:125-153.

Peterson, R. A., and S. P. Brown. 2005. On the use of beta coefficients in meta-analysis. Journal of Applied Psychology **90**:175.

Pietschnig, J., M. Siegel, J. S. N. Eder, and G. Gittler. 2019. Effect declines are systematic, strong, and ubiquitous: A meta-meta-analysis of the decline effect in intelligence research. Frontiers in Psychology **10**:2874.

Rodgers, M. A., and J. E. Pustejovsky. 2021. Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. Psychological methods **26**:141.

Rose, A. K., and T. D. Stanley. 2005. A meta‐analysis of the effect of common currencies on international trade. Journal of economic surveys **19**:347-365.

Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. Methods in Ecology and Evolution **1**:103-113.

Schmidt, F. L., and I.-S. Oh. 2013. Methods for second order meta-analysis and illustrative applications. Organizational Behavior and Human Decision Processes **121**:204-218.

Senior, A. M., C. E. Grueber, T. Kamiya, M. Lagisz, K. O'dwyer, E. S. Santos, and S. Nakagawa. 2016. Heterogeneity in ecological and evolutionary meta - analyses: its magnitude and implications. Ecology **97**:3293-3299.

Van Klink, R., D. E. Bowler, K. B. Gongalsky, A. B. Swengel, A. Gentile, and J. M. Chase. 2020. Meta-analysis reveals declines in terrestrial but increases in freshwater insect abundances. Science **368**:417-420.

Viechtbauer, W. 2010. Conducting meta-analyses in R with the metafor package. Journal of statistical software **36**:1-48.

Wasserstein, R. L., and N. A. Lazar. 2016. The ASA statement on p-values: context, process, and purpose. The American Statistician **70**:129-133.

Yang, Y., H. Hillebrand, M. Lagisz, I. Cleasby, and S. Nakagawa. 2022. Low statistical power and overestimated anthropogenic impacts, exacerbated by publication bias, dominate field studies in global change biology. Global Change Biology **28**:969-989.

FIGURE LEGENDS

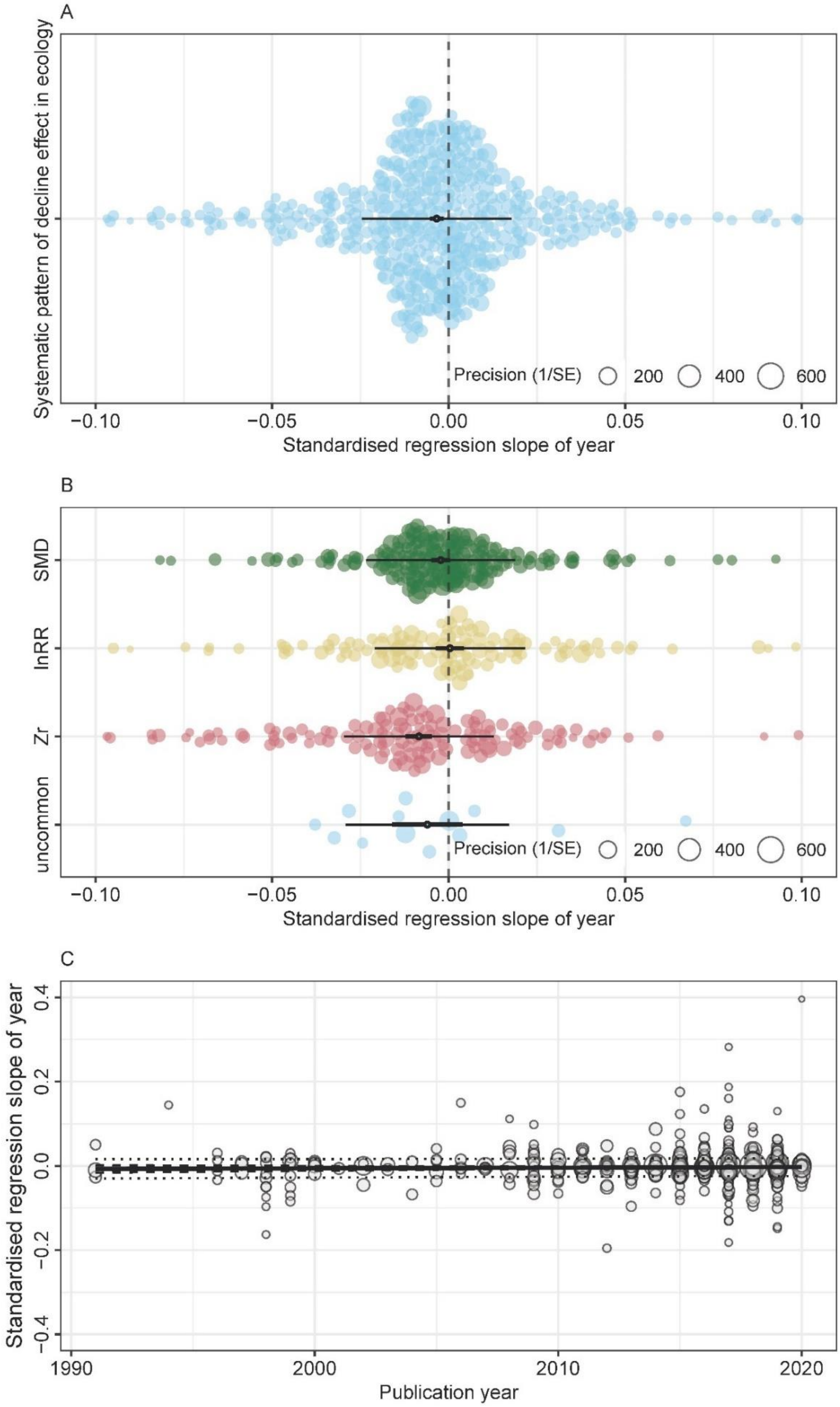

Figure 1. Orchard plots showing the distribution of regression slopes of year (β_{year}; solid circles) obtained from the first step (decline effect tests for all individual meta-analyses) and meta-analytic aggregation of β_{year} in the second step (mean β_{year}; open circle). (A) Mean β_{year} over individual β_{year}, which was used to identify the systematic pattern of decline effect. (B) Mean β_{year} for each effect size metric. (C) The relationship between meta-analyses' decline effect and the publication year (the test of the "decline in the decline effect itself"). The size of each point (individual β_{year}) is proportional to its precision (inverse standard error [SE] of β_{year}). Thick error bars represent 95% confidence intervals (CIs) and thin error bars represent prediction intervals (PIs). 95% CIs that do not cross zero (p-value < 0.05) indicate a "global" decline effect in ecology. lnRR = log response ratio; standardized mean difference = SMD, *Zr* Fisher's r-to-z, uncommon effect sizes = uncommonly used ecological effect size metrics (e.g., odds ratio). All plots were made using *orchaRd* package (Nakagawa et al. 2021).

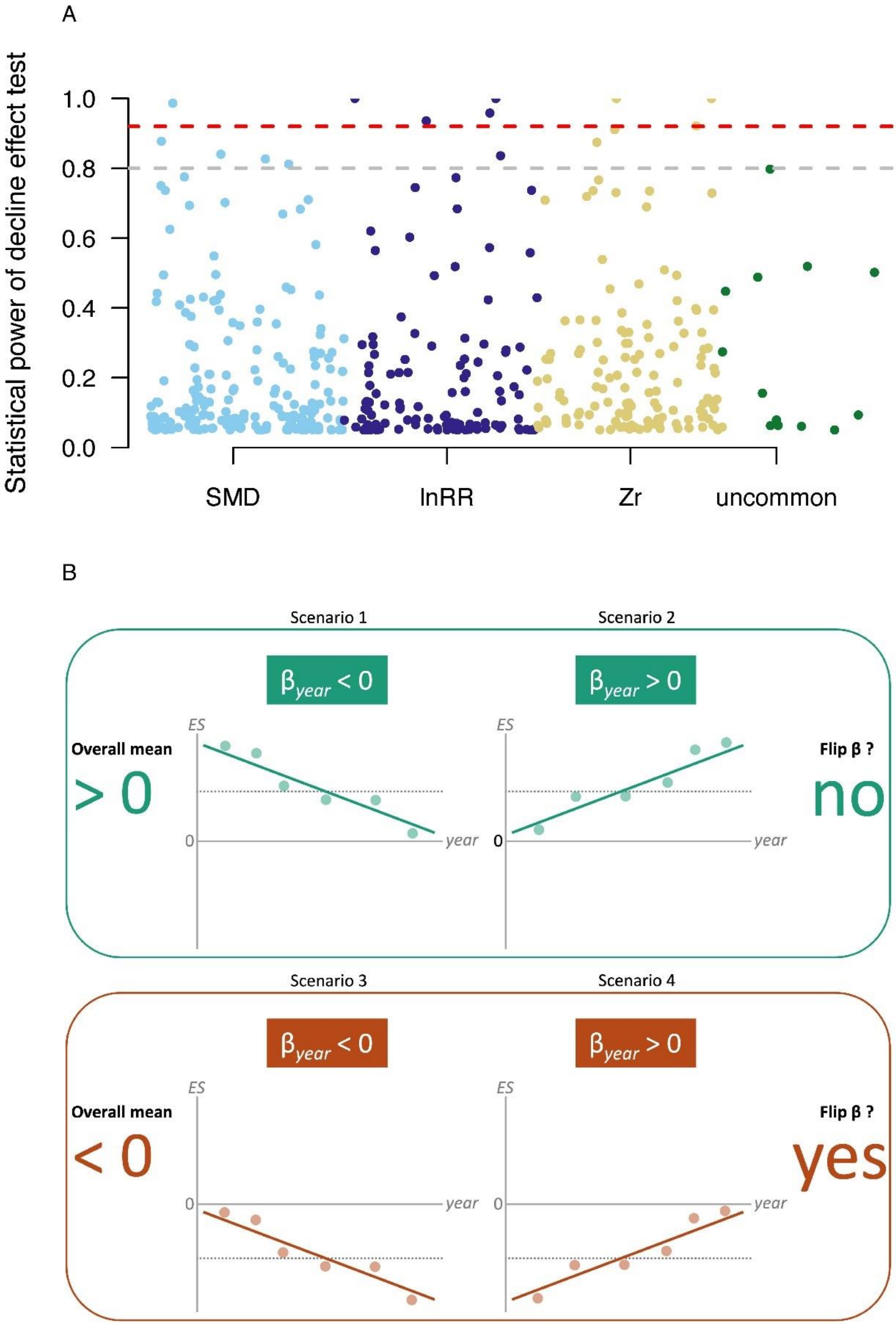

Figure 2. Statistical power of decline effect tests in individual meta-analyses (A) and justification of flipping the signs of β_{year} when detecting the systematic pattern of the decline effect (B). (A) The scatter plot of statistical power of decline effect test for all individual meta-

analyses from Costello and Fox (2022)'s dataset. The grey dashed line represents the nominal power level (i.e., 80%), below which the power is insufficient to detect a decline effect. The red dashed line represents the statistical power of the decline effect test performed using meta-meta-analysis (92%). (B) Different scenarios of flipping the signs of β_{year}. If a meta-analysis's mean effect size (overall mean) is positive, a positive β_{year} indicates that the magnitude of the effect size estimate declines over time (Scenarios 1 & 2). In contrast, for a meta-analysis with a negative mean effect size, a negative β_{year} indicates that the magnitude of the effect size estimate declines over time (Scenarios 3 & 4).