**BBSRC 2019 Button and Sena Practical**

**Assessing methodological quality of experimental studies**

We are designing a randomised controlled trial (RCT), comparing results-free review with review-as-usual on reducing publication bias. The aim of the trial is to evaluate a peer-review process, whereby editors and reviewers are blinded to the results, testing the impact this may have on reducing pre-existing bias towards the publication of ‘positive’ results obtained from low-quality studies.

The aim of this practical is to pilot the primary outcome measure developed to assess methodological rigour of experimental studies spanning biomedicine.

Participants will work in pairs to rate the methodological quality of a range experimental studies spanning biomedicine using a new outcome evaluation platform currently under development for use in the RCT. The goals are (i) for students to gain an understanding of the core principles of good experimental design that span disciplines (ii) to participate in the development of a robust generic measure for assessing methods rigour.

A brief protocol of the RCT is provided below.

**A randomised controlled trial to evaluate Results-Free Peer Review (RFPR) to reduce publication bias: a brief protocol for BBSRC 2019**

**6th January 2019 version 1.1**

## Background and rationale

Publication bias is the term for what occurs whenever the research findings in the published literature differ systematically from of the population of all studies completed in a given area ([Rothstein et al., 2005](#_17dp8vu)). Publication bias is wasteful, hindering science’s incremental progress ([Ioannidis, 2012](#_4d34og8)). Publication bias arises from the decisions of investigators, reviewers, and editors to submit or accept manuscripts for publication based on certain study characteristics. This could arguably be beneficial if decisions were made solely on study quality ([Dickersin, 1990](#_30j0zll)). The purpose of scientific enquiry is to estimate the presence and size of causal associations, and results from studies with the highest methodological quality will be the most reliable and informative. However, publication bias arises because of decisions influenced by the direction or strength of the study findings.

There is a long tradition in the study of publication bias ([Sterling, 1959](#_35nkun2); [Shadish et al., 1989](#_3rdcrjn); [Smart, 1964](#_26in1rg); [Greenwald, 1975](#_3dy6vkm); [Coursol and Wagner, 1986](#_gjdgxs); [Mahoney, 1977](#_2s8eyo1); [Smith, 1980](#_lnxbz9)). As early as 1959, Sterling found that of all the articles which used tests of significance published in 4 journals, 97% found in favour of the alternative hypothesis. Despite early awareness of publication bias, many disciplines have been slow to address the issue. While negative results are seemly disappearing from the literature across the sciences ([Fanelli, 2012](#_3znysh7)), papers testing hypotheses in psychology / psychiatry were found to show the greatest degree of bias for positive findings with over 90% of papers rejecting the null hypothesis ([Fanelli, 2010](#_1fob9te)).

Evidence for the causes of publication bias suggests that they occur throughout the publication process, from author decisions about what to write up through to reviewer recommendations and editorial decisions about what to publish. A study analysing a discrete population of conducted studies (Time-Sharing Experiments in the Social Sciences, k = 221) found that strong results were 40% more likely to be published, and 60% more likely to be written up, than null results ([Franco et al., 2014](#_tyjcwt)). When asked why they choose not to write up their null findings, 15 out of the 26 authors who replied suggested it was in the belief that null results have little publication potential ([Franco et al., 2014](#_tyjcwt)). This is consistent with evidence form clinical trials where positive trials were submitted for publication significantly more rapidly after completion than negative trials (median, 1.0 vs 1.6 years; P=.001) ([Ioannidis, 1998](#_1t3h5sf)).

Author decisions not to write-up null findings are likely based on a valid appreciation of the uphill struggle that null papers face during the review process. Reviewers have been shown to be highly influenced by the direction and strength of effects ([Dickersin, 1990](#_30j0zll)). On average null papers take several months longer from the time of submission to eventual publication than positive papers (median, 1.1 vs 0.8 years; P=.04, suggesting that null results have a harder time during the peer review process (Ioannidis, 1998). Indeed, the methods and quality of null studies are more critically reviewed than those of positive studies. An experimental study ([Mahoney, 1977](#_2s8eyo1)) randomly assigned referees to review 1 of 5 versions of a manuscript, all with identical introduction and methods sections, but different results and discussion sections (positive, negative, methods only, mixed results with positive discussion, mixed results with negative discussion). The methods, data presentation, scientific contribution, and publication merit of manuscripts with positive results were rated as being nearly twice as high as manuscripts with negative results. Interestingly, the scientific contribution and publication merit ratings for the manuscripts where reviews were based solely on the introduction and methods, were comparable to those for the positive results. Thus, negative findings seem to disproportionally and detrimentally affect appraisals of study quality and merit.

One way to prevent publication bias in the review and editorial process, is to base the decision to publish solely on assessing the scientific merit of the study rationale, and methods without access to the results. During ‘results-free’ review, authors submit other-wise complete manuscripts, but omit any discussion of results, and provisional accepted is based on peer-review of the introduction and methods alone. The results and discussion of accepted manuscripts are then reviewed in second stage, to check for adherence to methods and to allow minor revisions.

Attempts to evaluate the effectiveness of publishing initiatives are often confounded by a form of self-selection bias; authors who are invested in the principles of scientific rigour and transparency are more likely to submit their work to special issues piloting results-free review. Similarly, such initiatives may attract reviewers that are more switched on to the research integrity and publication bias debate.

To overcome these limitations, the we propose a randomised controlled trial (RCT), comparing results-free review with review-as-usual on reducing publication bias.

## Objectives

To evaluate a peer-review process, whereby editors and reviewers are blinded to the results, testing the impact this may have on reducing pre-existing bias towards the publication of ‘positive’ results obtained from low-quality studies.

## Trial design

A parallel arm randomised controlled trial (RCT), with 1:1 randomisation.

1.1. Aim: To evaluate whether a 2-stage review process whereby the decision to accept or reject the manuscript is based on an initial review of a results-free version of the manuscript, reduces publication bias in terms of raising the methodological quality of published articles, and reducing the proportion of 'positive' primary findings.

1.2. Population: Manuscripts submitted to BMC Psychology and other journal in the Nature-Springer group recruited to the study which have been determined in the initial screening phase to report experimental research.

1.3. Intervention: Results-Free Peer Review (RFPR). The results and discussion are removed from manuscripts before being sent to editors and reviewers who base their decision to accept/reject on review of the rationale and methods alone. The results and discussion of accepted manuscripts are then reviewed in second stage, to check for adherence to methods and to allow minor revisions.

1.4. Comparator: Review-as-usual, manuscripts managed according to current standard practices.

1.5. Outcomes:

1) Difference in ratings of methodological quality of research reported in manuscripts accepted for publication under RFPR and review-as-usual.

2) Difference in proportion of accepted manuscripts reporting 'positive' primary results in RFPR versus review-as-usual.

# Methods

## Study setting

BMC Series and other Nature-Springer journals recruited into the study.

## Eligibility criteria

All original articles describing *in vivo* or *in vitro* experimental research using humans or non-human animals, will be eligible for inclusion.

***Exclusion criteria***

Observational epidemiology studies. Secondary analyses of existing datasets. Systematic reviews and meta-analysis of existing publications.

## Intervention

Manuscripts will be randomised to intervention or control. Those randomised to RFPR (intervention) will have their results and discussion sections redacted from the abstract and main text body by the in-house technician. Those manuscripts randomised to the review-as-usual (control) will remain whole.

In the intervention arm the editorial decisions are made in two-stages. The first stage is ‘results-free’ and if the methods and rationale are deemed sound, the Associate Editor offers an ‘acceptance-in-principle’. If not, the manuscript is rejected. The second stage follows an ‘acceptance-in-principle’. For those manuscripts accepted-in-principle the results and discussion are unredacted and the original reviewers are asked to review the full manuscript, inclusive of results and discussion to check the adherence to the methods and that the inferences drawn are supported by the data. Revisions may be suggested, but the decision to publish will only be revoked in cases where the results and discussion deviate extensively and unjustifiably from the stated aims and methods.

## Outcome assessment

Each accepted manuscript will be scored by two independent reviewers blinded both to intervention status and to the other reviewer's scores using a tool developed for this purpose, enabled on an online platform. Discrepancies will be resolved by a third reviewer, who will be blinded to the identity of the previous two reviewers but unblinded to their scores.

For the primary outcome, we will create a checklist to assess the rigour of study design based on an operationalisation of the relevant items from the following reporting guidelines (Landis et al 2012) Core Set of Reporting Standards for Rigorous Study Design, with additional items from ARRIVE and CONSORT reporting guidelines and Cochrane Risk of Bias Tool.

For the secondary outcome in line with (Fanelli, 2010) papers will be classified as ‘positive’ if they report full or partial support for the primary tested hypothesis or research question in the abstract, or ‘negative’ if they find no or null support for the tested primary hypothesis as stated by the authors in the abstract.

Individuals experienced in the critical appraisal of published research materials will be recruited, and will be required to complete a training program on the web platform before beginning the project. The training program will consist of 10 papers with gold standard answers (set by the research team) of which the individual will need to score 80% on minimum of 3 consecutive papers to pass, and pass all of the critical items (randomisation, blinding, sample size justification, data handling). The qualified reviewers will be randomly allocated the manuscripts. Each manuscript will be assigned to two reviewers. The online system will ensure that reviewers do not receive, for second review, a manuscript for which they performed the first review.

***Primary outcome***

* The difference in the proportion on the rigour checklist items met for each study between the intervention and control arms.

***Secondary outcome***

* The difference in the proportion of manuscripts reporting positive primary results between the intervention and control arms.

# References

Coursol A and Wagner EE. (1986) Effect of positive findings on submission and acceptance: A note of meta-analysis bias. *Professional Psychology: Research and Parctice* 17: 136-137.

Dickersin K. (1990) The existence of publication bias and risk factors for its occurrence. *JAMA* 263: 1385-1389.

Fanelli D. (2010) "Positive" results increase down the Hierarchy of the Sciences. *PLoS One* 5: e10068.

Fanelli D. (2012) Negative results are disappearing from most disciplines and countries. *Scientometrics*: 891–904.

Findley MG, Jensen NM, Malesky EJ, et al. (2016) Can results-free review reduce publication bias? The results and implications of a pilot study *Comparative Political Studies*: 1-37.

Franco A, Malhotra N and Simonovits G. (2014) Publication bias in the social sciences: Unlocking the file drawer. *Science* 345: 1502-1505.

Greenwald AG. (1975) Consequences of prejudice against the null hyptohesis *Psychol Bull* 82: 1-20.

Ioannidis JP. (1998) Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 279: 281-286.

Ioannidis JPA. (2012) Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science* 7: 645-654.

Mahoney MJ. (1977) Publication prejudices: An experimental study of confirmatory bias in the peer review system *Cognit Ther Res* 1: 161-175.

Rothstein H, Sutton AJ and Borenstein M. (2005) *Publication bias in meta-analysis : prevention, assessment and adjustments,* Chichester: John Wiley.

Shadish WR, Jr., Doherty M and Montgomery LM. (1989) How many studies are in the file drawer? An estimate from the family/marital psychotherapy literature *Clin Psychol Rev* 9: 589-603.

Smart RG. (1964) The importance of negative results in psychological research. *The Canadian Psychologist* 5: 225-232.

Smith ML. (1980) Sex bias in counseling and psychotherapy *Psychol Bull* 87: 392-407.

Sterling TD. (1959) Publication Decisions and Their Possible Effects on Inferences Drawn from Tests of Significance--Or Vice Versa. *Journal of the American Statistical Association* 54: 30-34.