Dealing with publication bias in translational stroke research

Shimin Liu, M.D., Ph.D.*
Department of Neurology, Mount Sinai School of Medicine, NYU

Abstract
Publication bias has been around for about 50 years. It has become a concern for almost 20 years in the medical research community. This review briefly summarizes the current status of publication bias, potential sources where bias may arise from, and its common evaluation methods. In the field of translational stroke research, publication bias has long been suspected; however, it has not been addressed with sufficient efforts. Its status has remained the same during the last decade. The author emphasizes the important role that publishers might play in addressing publication bias.

Keywords: Publication bias, stroke, funnel plot, preclinical trials, neuroprotection, efficacy

Correspondence should be sent to:
Dr. Shimin Liu, M.D., Ph.D.,
Department of Neurology, Mount Sinai School of Medicine NYU.
1468 Madison Avenue, New York, NY 10029
E-mail: shimin.liu@mssm.edu

1. Introduction
Publication bias has been noticed for about 50 years (Sterling 1959), and has become a concern for almost 20 years (Chalmers et al 1990; Dickersin 1990; Sharp 1990). In most cases, it refers to the assertion that studies having positive and/or statistically significant results are more easily and frequently published. Other types of publication bias may include a preference to publish based on research directions, authors' nationalities, and institutes' professional ranks. Publication bias in the medical research reporting system may have distorted the medical research literature and may have influenced the conclusion of some meta-analysis (systematic) reviews. A small scale random sample of 86 studies (Machan et al 2006) in medical informatics evaluation research showed a remarkably high percentage (69.8%) of descriptions of positive results. 19 (36.6%) of the analyzed 54 reviews and meta-analyses came to a positive conclusion with regard to the overall effect of the analyzed system, 32 (62.5%) were inconclusive, and only one review came to a negative conclusion.

2. Publication bias in translational stroke research
In the field of translational stroke research, the discrepancy of the neuroprotective efficacy between preclinical trials and clinical trials has caused growing concerns. An extensive review of 1026 experimental treatments (O'Collins et al 2006) revealed that neuroprotective efficacy was superior to control conditions in 62% of the preclinical models of focal ischemia, in 70% of preclinical models of global ischemia, and in 74% of culture models. Such a rate of reporting positive results in preclinical trials is drastically high when compared with the rate of reporting positive results in clinical trials. Currently we still have no FDA-approved neuroprotective treatment for ischemic stroke.

Factors contributing to the translational failures of neuroprotective treatments for ischemic stroke have been addressed in the Stroke Therapy Academic Industry Roundtable (STAIR) guidelines for preclinical stroke trials (Stroke Therapy Academic Industry Roundtable 1999), which include species differences, inappropriate time windows, ineffective drug levels, inability of drugs to cross the blood-brain-barrier (BBB), use of young animals without co-morbidities, failure to model white matter damage, and the heterogeneity of stroke subtypes in patients. However, another issue, the known publication bias, may also play a role in causing the discrepancy of positive result reporting between preclinical trials and clinical trials in translational stroke research.

An article from Collaborative Approach to Meta Analysis and Review of Animal Data from Experimental Stroke (CAMARADES) dealing with publication bias across stroke studies suggests that at least 15% of experiments remain unpublished, and that this results in an overstatement of efficacy of 30% (unpublished communication with Dr. Macleod).

Some reviews with detailed systematic analyses further confirmed the widely suspected publication bias in preclinical stroke trials (Perel et al 2007). As manifested on the funnel plot, which is the most commonly used graphical evaluation for publication bias, an inflated efficacy was shown with the treatment of re-
combinant tissue plasminogen activator, and similar effects were shown with tirilazad treatments (Perel et al 2007), but not with hypothermia treatment (van der Worp et al 2007).

3. Sources of bias

Although it is widely noticed that positive or significant results more frequently appear in journals, the fundamental reasons have not been well addressed. The bias that has been noticed in medical research literature might come from the publication process, the experimental process, or both. Currently there is no particular study specifically addressing the bias sources in translational stroke research. It is reasonable to assume that similar bias sources exist in translational stroke research as in other medical research fields.

3.1. Bias from the publishing process

In addition to the result being positive or negative (Blackwell et al 2009; Hopewell et al 2009), many other factors may contribute to publication bias, such as research directions of authors and reviewers (Joyce et al 1998), manuscript’s potential value for pursuing and maintaining a journal’s high impact factor (Opthof et al 2002b), conflicts of interest (Perlis et al 2005), research funding sources (Liss 2006), publishing cost, language (Egger et al 1997b), author’s nationality (Opthof et al 2002b; Yousefi-Nooraie et al 2006), the rank and geographic location of the sponsor institute (Eloubeidi et al 2001; Sood et al 2007), and multiple biases from the review process (Alasbali et al 2009; Goldbeck-Wood 1999; Opthof et al 2001a; Opthof et al 2002a; Opthof et al 2002b). The following paragraphs mainly discuss the influence of sponsorship and impact factor on publication bias.

It has been noticed that research funding sources influence the published outcomes of studies. Results favorable for the drugs studied were significantly more common in those funded by a pharmaceutical company (98% vs. 32%) (Liss 2006). Financial conflicts of interest have been reported to be prevalent in clinical trials and are associated with a greater likelihood of reporting results favorable to the intervention being studied (Perlis et al 2005). Research funded by drug companies was less likely to be published than research funded by other sources (Hall et al 2007; Lexchin et al 2003). Studies sponsored by pharmaceutical companies were more likely to have outcomes favoring the sponsor than were studies with other sponsors (Lexchin et al 2003). Therefore, additional procedures may need to be taken for avoiding bias as well as a declaration of conflict of interest.

While the impact of research funding sources on publication bias has been noticed and addressed, the role of a journal’s operating goal and supporting sources has rarely been discussed. Although it is controversial for using the impact factor as a criterion for measuring a journal’s quality (Barendse 2007; Boldt et al 2000; Peleg and Shvartzman 2006; Rous-sakis et al 2007), most journals tend to treat the impact factor as a measure of their journal’s achievements. Publisher bias may be encouraged when the impact factor prevails in a journal’s operational strategy (Opthof et al 2001b; Opthof et al 2002b). A study assessed the relationship of a journal’s impact factor and publishing outcomes in the literature of neonatology (Litnner et al 2005). It showed that studies with positive results were more frequently published in journals with high impact factors, suggesting a role of the impact factor in selective publishing or submitting. Seeking a different operating strategy for professional journals may be needed for dealing with publication bias.

3.2. Bias from experimental process

Bias coming from the experimental process may mix with the publication bias; it may not be easy to tell whether the bias is from the publication process, the experimental process, or both. Experimental bias has been well-noticed and addressed in a few systematic reviews (Macleod et al 2008b; van der Worp et al 2007). Some editors may argue that researchers are more likely to produce and recommend manuscripts with positive or significant results. JAMA did a specific investigation on publication bias in editorial decision making (Olson et al 2002); after having adjusted simultaneously for study characteristics and quality indicators, the publication rates between studies with significant and non-significant results did not differ significantly, with an adjusted odds ratio of 1.30 (95% CI, 0.87-1.96). However, editorial policies and processes differ from journal to journal, and a study based on one journal’s articles is insufficient to reflect the overall status of journals’ roles in publication bias. JAMA are to be praised for their constructive efforts in addressing the issue of publication bias from the journal’s side.

4. Evaluation methods for publication bias

There is no specially designed or tailored evaluation method for the detection of publication bias in translational stroke research. To estimate a suspected publication bias in stroke studies, all regular methods that are widely accepted by the medical research community will apply. Publication bias can be detected by several commonly used graphic, or statistical, methods, such as the funnel plot (Egger et al 1997a) or fail safe numbers (Persaud 1996; Rosenberg 2005). Other methods, such as selection models using weighted distribution theory (Sutton et al 2000), are also in development, but they have not been used...
regularly. Brief introductions for the common evaluation methods and sample funnel plots for the efficacies of some neuroprotectants are provided in the following paragraphs.

The funnel plot, the plot of a trial’s effect estimates against sample size, has been widely used to deal with publication bias. It detects bias based on the assumption that the plot resembles a symmetrical inverted funnel in the absence of bias (Egger et al 1997a). Many factors may potentially contribute to the detected asymmetry; therefore, this method should be used with caution, especially when limited numbers of studies are used in a meta-analysis (Irwig et al 1998; Stuck et al 1998; Vandenbroucke 1998).

Detection of asymmetry in a funnel plot can be conducted by several methods, such as visual inspection, “trim and fill”, regression approach (Soeken and Srupusanpan 2003), and a newly emerged method (Formann 2008) in which the proportion of unpublished studies is estimated by the degree of truncation from a left-truncated normal distribution. Each method has its own advantages and limitations; it is suggested that multiple methods should be used when there is a suspicion of publication bias.

For demonstration purposes, here I present a sample funnel plot basing on the available data from a systematic review paper (Macleod et al 2005a) for the neuroprotective efficacy of FK506. As shown in Fig. 1, the funnel plot of precision against effective size is asymmetric by visual inspection, with more studies allocated in the left side. The “missing” studies in the right portion (large efficacy) of the plot may indicate possible publication bias or true heterogeneity of studies. A similar sample funnel plot (Fig. 2) for the neuroprotective efficacy of melatonin has also been completed based on the available data from another systematic review (Macleod et al 2005b). Being different from Fig 1 by visual inspection, the asymmetry of Fig 2 reveals apparent “missing” studies in the left portion of small efficacy, indicating a possibly inflated efficacy. These asymmetries in funnel plots may indicate the existence of publication bias, missing literature, or true heterogeneity among studies. Therefore, addressing suspected publication bias in systematic reviews may provide valuable information for the selection of neuroprotective candidates for acute stroke treatment.

Fail safe numbers are to some degree analogous to the confidence intervals. They aid in the assessment of the degree of confidence for a particular result in meta-analysis studies, which the funnel plots do visually. Fail safe numbers may be defined as the number of new, unpublished, non-significant studies that would be required to exist to lower the significance of a meta-analysis to some specified level (Persaud 1996). Using fail-safe numbers is a quick way to estimate whether publication bias is likely to be a
problem for a specific meta-analysis study (Rosenberg 2005).

Some graphic evaluation methods frequently appear in systematic reviews, such as the forest plot, the Galbraith plot, the L’Abbé’ plot, and the box plot; but they are designed for assessing study variation and heterogeneity, not specifically for assessing publication bias (Bax et al 2008).

5. Multiple initiatives suggested for reducing publication bias

About 50 years have passed since the notice of publication bias (Sterling 1959), yet little has changed. A significant amount of research conducted never appears in the public forum. It has been estimated that less than half of all studies initially presented as summaries or abstracts at professional meetings are subsequently published as peer-reviewed fulltext articles (Scherer et al 1994; Scherer et al 2007), although various factors may account for such a low publication rate.

Publication bias influences the conclusion of systematic reviews; therefore it should be addressed appropriately when there is a suspicion. However, the methods used for dealing with publication bias in most systematic reviews appear merely as a description of the number of pooled studies, the number of searched databases, or the statistical analysis of heterogeneity. In attempts to address this situation, both the general medical research community and the translational stroke research community have started multiple initiatives.

5.1. Initiatives in the general medical research community

Some journals have implemented guidelines to make sure that the issue of publication bias be dealt with, such as the QUOROM statement, and MOOSE guidelines. The quality of reports of meta-analyses (QUOROM) statement was suggested for meta-analyses of randomized controlled trials (Moher et al 2000a; Moher et al 2000b). The meta-analysis of observational studies in epidemiology (MOOSE) guidelines, the MOOSE checklist, was suggested for meta-analyses of observational studies (Stroup et al 2000). Some journals implement a general requirement to address publication bias using an appropriate method when necessary. More detailed descriptions of dealing with publication bias have been provided in the Cochrane handbook for Systematic Reviews of Interventions.

In an effort to balance publication bias in medical literature, Journal of Negative Results in Biomedicine (http://www.jnrbm.com/) was started in 2004 to accept manuscripts with unexpected, controversial, provocative and/or negative results/conclusions.

Although not being established particularly for addressing publication bias, ClinicalTrials.gov (http://clinicaltrials.gov/) provides a global view of clinical trials. ClinicalTrials.gov is a registry of federally and privately supported clinical trials conducted in the United States and around the world. ClinicalTrials.gov currently contains 66,791 trials sponsored by the National Institutes of Health, other federal agencies, and private industry. Studies listed in the database are conducted in all 50 States and in 161 countries. A similar registration agent for preclinical trials will be expected to provide a global view for preclinical trials, to counter against publication bias, and to help with the selection of candidates for clinical trials.

5.2. Initiatives in the translational stroke research community

As in the meta-analyses of cardiovascular diseases (Palma and Delgado-Rodriguez 2005), the assessment of publication bias in systematic reviews of preclinical and clinical trials for ischemic stroke treatment appears at a low frequency. Although methodology details and recommendations for assessing publication bias have been provided in the Cochrane handbook for Systematic Reviews of Interventions (http://www.cochrane-handbook.org/), none of the Cochrane’s 35 systematic reviews (http://www.cochrane.org/reviews/en/subtopics/93.html) on medical therapies for ischemic stroke used the widely accepted methods, such as funnel plots, or failsafe-N. Only the heterogeneity of studies has been assessed in these reviews. It is a good sign to see that a few other reviews have addressed publication bias in professional details (Perel et al 2007; van der Worp et al 2007).

Because the bias that appears in published materials may also come from the experimental process, efforts in reducing experimental bias may help to reduce the observed bias in the literature. More details about reducing experimental bias have been described in Good Lab Practice Guidelines (GLPG) (Macleod et al 2008a). Measures such as random allocation, random sampling, blind assessment, and standardized operational procedures could be used for reducing experimental bias. This GLPG, together with the STAIR criteria, if being followed strictly, may help to reduce the inflated efficacy in preclinical trials for the treatment of acute ischemic stroke. However, almost ten years have passed since the first version of the STAIR criteria (Stroke Therapy Academic Industry Roundtable 1999), yet the efficacy discrepancy between preclinical and clinical stroke trials doesn’t seem to have improved. An independent registration and validating system may help to reduce the bias from both publication process and experimental process. Seeking a different approach is ne-
cessary to reduce the impact of this problem (Hall et al. 2007).

The CAMARADES collaboration has been active in addressing publication bias in experimental stroke. Some of their focuses include identifying potential sources of bias in animal work, developing recommendations for improvements in the design and reporting of animal studies, and developing better meta-analysis methodologies for animal studies. The CAMARADES group will soon launch an on-line facility for the registration of animal studies in stroke with enough details to help systematic reviewers contact authors of unpublished studies. This would be a good start for the preclinical stroke trial registration, and more work will still be needed to establish a world-wide registry. Research sponsors and governmental authorities may be suggested as needing to become involved in promoting mandatory registration.

6. How JESTM addresses publication bias?

The goal of the Journal of Experimental Stroke & Translational Medicine (JESTM, www.jestm.com) is to foster new concepts and to reflect the status of preclinical trials in the field of experimental and translational stroke research. Therefore, manuscripts with controversial/provocative ideas and negative results are encouraged equally.

In order to increase the quality of our published contents and to reduce publication bias, we require authors of review manuscripts be aware of and to address publication bias appropriately. We require authors of research studies to conform to Good Lab Practice Guidelines for reducing experimental bias.

JESTM is operated with support from enthusiastic professional volunteer workers. Because it is purely an online journal, its operational cost is considerably lowered. All articles on JESTM are free-access and authors are not charged for publishing cost. Articles will be made immediately available online after completion of the requisite review process.

Acknowledgement

This work was supported by NIH grant 5T32NS051147-02 and NS 21076-24. The author appreciates and acknowledges Dr. Levine and Dr. Winn at Mount Sinai School of Medicine for his contribution on revising this paper.

References


Macleod MR, O’Collins T, Horky LL, Howells DW, Donnan GA (2005a) Systematic review and metaanalysis of the efficacy of


Sharp DW (1990) What can and should be done to reduce publication bias? The perspective of an editor. *Jama* 263:1390-1.


Stuck AE, Rubenstein LZ, Wieland D (1998) Bias in meta-analysis detected by a simple, graphical test. Asymmetry detected in funnel plot was probably due to true heterogeneity. *Bmj* 316:469; author reply 70-1.


