Chapter 7

The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: the case of depression

Ymkje Anna de Vries, Annelieke M. Roest, Peter de Jonge,
Pim Cuijpers, Marcus R. Munafò, Jojanneke A. Bastiaansen

Submitted

Abstract

Study publication bias is recognized as an important threat to evidence-based medicine, but other biases also affect the quality of the evidence base. Using the evidence base for antidepressants and psychotherapy, we illustrate how the effects of study publication bias, outcome reporting bias, spin, and citation bias accumulate to hide negative results from view.

Within the antidepressant literature, 52 (50%) of 105 trials were negative, but only 4 (5%) published trials unambiguously reported that the treatment was not effective. Other negative trials remained unpublished (27 trials), were published as if positive due to outcome reporting bias (10 trials), or were published with spin in the abstract or discussion (11 trials). Compounding the problem, trials reporting positive results were cited, on average, three times as frequently as negative trials (92 versus 32 citations).

Within the psychotherapy literature, we obtained 142 published trials. Although 49 (35%) of these trials were considered to be negative, only 12 (8%) abstracts concluded that psychotherapy was not more effective than a control condition. Positive psychotherapy trials were also cited more frequently than negative trials (111 citations versus 58).

These results show the pernicious cumulative effect of reporting and citation biases. Even when the decision to publish a trial is made, there are still several hurdles to pass before negative results can receive the visibility they deserve. Mandatory universal registration, in combination with openness to negative results and vigilance on the part of peer reviewers, journal editors, and readers, may help to prevent and uncover bias.

Introduction

Evidence-based medicine is the cornerstone of good clinical practice, but it is dependent on the quality of the evidence upon which it is based [323]. It is well-known that trials with statistically significant findings are more likely to be published than those with non-significant findings; this is a major problem since as many as half of all randomized controlled trials (RCTs) are never published [38].

However, negative trials face additional hurdles beyond study publication bias that can result in the disappearance of non-significant results [38, 40, 44]. Here, we analyze the cumulative impact of biases on the apparent efficacy of treatments, and discuss possible remedies, using the evidence base for two effective treatments for depression: antidepressants and psychotherapy [19, 324].

Reporting and citation biases

Many different biases exist and can affect an evidence base. We distinguish between four major biases: study publication bias, outcome reporting bias, spin, and citation bias. While study publication bias involves non-publication of an entire study, outcome reporting bias refers to non-publication of negative outcomes or non-significant analyses within a published article [38]. It also refers to relegating non-significant primary outcomes to a secondary status or upgrading a protocol-specified secondary (but statistically significant) outcome to a primary outcome in the publication. Both study publication bias and outcome reporting bias can be an important threat to the validity of meta-analytic conclusions regarding treatment efficacy [102, 325].

Although trials that faithfully report non-significant results on the primary outcome will yield accurate effect estimates for use in meta-analyses, the interpretation of results can still be positively biased, which may affect the apparent efficacy of treatments. The use of specific reporting strategies that could distort the interpretation of results and mislead readers is defined as spin [40]. Spin occurs, for instance, when authors focus on statistically significant results in secondary analyses and conclude that the treatment is effective despite non-significant results on the primary outcome. It has been shown that a treatment is rated as being more beneficial when the abstract of a journal article has been spun [41].

Finally, even when a trial with non-significant results is published and accurately reported, citation bias is an additional obstacle to ensuring that it receives as much attention as a study with significant results. Studies with positive results receive more citations than negative studies in many medical fields, including psychiatry [45, 313], which results in a heightened visibility of positive results.

The evidence base for antidepressants

To examine the cumulative effect of these biases, we assembled a cohort of 105 trials of antidepressants for depression. Seventy-four of these trials were included in the 2008 study by Turner et al. [19], to which we added 31 trials from the Food and Drug Administration (FDA) database [105] used to support the marketing approvals of four novel antidepressants after the Turner et al. study. Pharmaceutical companies are required to preregister all trials that they intend to use in support of such an application with the FDA; hence, trials with non-significant results may not be published but are still accessible.

Figure 7.1 demonstrates the cumulative impact of reporting and citation biases on these trials. In the initial cohort of 105 antidepressant trials, 53 (50%) trials were considered to be positive by the FDA and 52 (50%) were considered negative or questionable (Figure 7.1a). While all but one of the positive trials was published (98%), only 25 (48%) of the negative trials were published. Hence, the initial cohort is reduced to 77 published trials, of which 25 (32%) were negative according to the FDA (Figure 7.1c).

Ten of these published negative trials, however, became 'positive' trials in the published literature, as a consequence of omitting unfavorable outcomes or switching the status of primary and secondary outcomes (Figure 7.1c). Without access to the FDA reviews, it would not have been possible to tell that these trials, when analyzed according to the pre-specified protocol, were not positive.

Among the remaining 15 (19%) negative trials, five were published (in four articles) with spin in the abstract (i.e., concluding that the treatment was effective). Five additional articles contained mild spin in the abstract (e.g., suggesting that the treatment is at least numerically better than placebo, or that the results cannot be interpreted because the trial, rather than the drug, failed, since an active comparator was not more effective than placebo either) [326]. One article did not have an abstract, but the discussion section concluded that there was a "trend for efficacy". Hence, only four (5%) of 77 published trials unambiguously reported that the treatment was not effective (Figure 7.1d).

Compounding the problem, trials reporting positive results were cited three times as frequently as negative trials (92 versus 32 citations in Web of Science, January 2016) (Figure 7.1e). Among negative trials, those with (mild) spin in the abstract received an average of 36 citations, while those with a clearly negative abstract received 25 citations on average. Although there were too few published negative antidepressant trials for a definitive conclusion, this suggests a possible synergistic effect between spin and citation bias, where negatively presented negative trials receive especially few citations [315, 327]. Altogether, these results show that the effects of different biases accumulate to hide non-significant results from view.

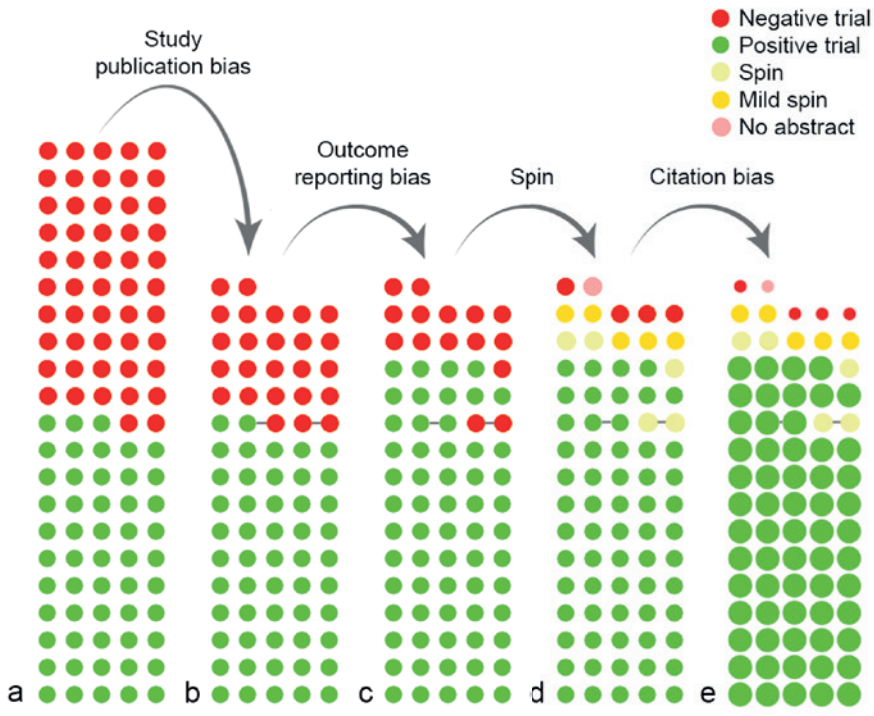

Figure 7.1: *Panel a displays the original, complete cohort of trials, while Panels b through e show the cumulative effect of biases. Each circle indicates a trial, while the color indicates the results (red for negative, green for positive) or the presence of spin (yellow and orange for strong and mild spin, respectively, and pink for a trial that was published in an article without an abstract). Circles connected by a grey line indicate trials that were published together in a pooled publication. In Panel e, the size of the circle indicates the (relative) number of citations received by that category of studies.*

The evidence base for psychotherapy

While the pharmaceutical industry has a financial motive for suppressing unfavorable results, these biases are also present in other areas of research, such as psychotherapy. In the absence of a standardized registry of trials, however, their presence is more difficult to assess, and the various biases are more difficult to disentangle. Statistical tests suggest an excess of positive findings in the psychotherapy literature, which may be due to either study publication bias or outcome reporting bias [177, 328].

More recently, in a cohort of psychotherapy trials funded by the National Institutes of Health, which maintains a database of funded grants and hence functions as a (limited) registry, it was found that 13 (24%) out of 55 initiated trials were never published [35]. The effect size of these unpublished studies was markedly lower than that of published

studies, suggesting a bias against publishing negative findings.

Regarding spin in the psychotherapy literature, we found that 49 (35%) of 142 papers were considered to be negative in a meta-analysis by Flint et al. [328], but only 12 (8%) abstracts concluded that psychotherapy was not more effective than a control condition. The remaining 130 abstracts reported either positive (73%) or mixed findings (19%), concluding, for example, that the treatment was effective for one outcome but not another.

Although we could not establish the pre-specified primary outcome for these trials, and therefore cannot determine whether a specific abstract is biased, it is clear that published psychotherapy trials, as a whole, provide an impression of the effectiveness of psychotherapy that is more positive than justified by the available evidence. Similar to the antidepressant literature, positive psychotherapy trials were cited more frequently than negative trials (111 citations versus 58). Compounding the problem, negative trials with a positive or mixed abstract received an average of 59 and 87 citations, respectively, while those with a negative abstract received only 26 citations.

Preventing bias

Mandatory and universal prospective registration has long been advocated as a solution to the widespread problem of study publication bias [329]. It could also be an effective solution to the problem of outcome reporting bias. Since 2005, the International Committee of Medical Journal Editors (ICMJE) has required prospective registration of clinical trials as a precondition for publication [330], but many journals do not require registration [331] and others allow retrospective registration [332]. Since 2007, investigators are also legally required by the FDA to prospectively register most phase 2 and 3 drug trials (i.e., clinical trials investigating efficacy and safety in patients).

The increasing pressure to publicly register trials may explain why negative trials of novel antidepressants are more frequently published than those of antidepressants approved earlier. In particular, all unpublished negative or questionable trials of novel antidepressants were completed before 2004, while the 25 trials completed in 2004 or later (including 14 trials for which registration was legally required) were all published, even though nine of these trials were considered negative or questionable.

A regulatory requirement is likely to be one of the most effective measures to ensure universal registration. It is therefore unfortunate that the 2007 law is not fully comprehensive, as it excludes trials of behavioral interventions (such as psychotherapy) and phase 1 (healthy volunteer) trials.

However, registration in and of itself also seems to be insufficient to ensure complete and accurate reporting of a trial. It has been found that only around half of all trials registered in ClinicalTrials.gov were published within two years of completion (as required by the

FDA Amendments Act) [333]. The COMPare Project (compare-trials.org), which aims to track outcome switching, and other studies [334] found that non-reporting of outcomes and the addition of novel outcomes (without any indication that these outcomes had not been pre-specified) was also common.

Hence, close examination of clinical trial registries by independent researchers may be necessary for trial registration to be a truly effective deterrent to study publication bias and outcome reporting bias. For already completed drug trials, consulting FDA reviews as we have done for this analysis may also be helpful. An alternative (or addition) to registration could be the publication of study protocols or "registered reports", in which journals accept a study for publication on the basis of the introduction and methods and before the results are known.

It may prove to be more difficult to prevent spin and citation bias. Peer reviewers could play a crucial role in ensuring that abstracts accurately report a trial's findings and that important negative results are cited. Journals also have a role to play in ensuring that researchers do not feel like they must 'oversell' the significance of their results to get published.

The prevalence of spin and citation bias also indicates the importance of assessing a study's actual results (rather than relying on the authors' conclusions) and of conducting an independent literature search whenever possible, since the reference lists of other papers may yield a disproportionate number of positive (and positively presented) studies.

Conclusions

The problem of study publication bias, which was first recognized over 50 years ago [36], is now well-known. Our examination of antidepressant trials, however, shows the pernicious cumulative effect of additional reporting and citation biases. Even when the decision to publish a trial has been made, there are still several hurdles to pass before negative results can receive as much visibility as positive results.

Within the antidepressant literature, we found that the cumulative impact of these biases eliminates the majority of negative results from the published literature and leaves the few remaining negative results that do get published much more difficult to discover than positive results. These biases are unlikely to be unique to antidepressant trials, which we have taken as a case study. We have shown that similar processes, though more difficult to assess, are at work within the psychotherapy literature.

The presence of these individual biases has been shown in various medical fields [38, 40, 313], and it is likely that their effects also accumulate whenever these biases are present. Consequently, researchers and clinicians across medical fields must be aware of the potential for bias to distort the apparent efficacy of an intervention.