ELSEVIER

Journal of Clinical Epidemiology ■ (2015) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Publication bias and small-study effects magnified effectiveness of antipsychotics but their relative ranking remained invariant

Dimitris Mavridis[a,b,*], Orestis Efthimiou[b], Stefan Leucht[c], Georgia Salanti[b]

[a]*Department of Primary Education, University of Ioannina, University Campus, PO Box 1186, 45110 Ioannina, Greece*
[b]*Department of Hygiene and Epidemiology, University of Ioannina, PO Box 1186, 45110 Ioannina, Greece*
[c]*Department of Psychiatry and Psychotherapy, Technische Universität München, Arcisstraβe 21, D-8033 Munich, Germany*

Accepted 29 May 2015; Published online xxxx

Abstract

Objectives: Publication bias (PB) may seriously compromise inferences from meta-analyses. The aim of this article was to assess the potential effect of small-study effects and PB on the recently estimated relative effectiveness and ranking of pharmacological treatments for schizophrenia.

Study Design and Setting: We used a recently published network of 167 trials involving 36,871 patients and comparing the effectiveness of 15 antipsychotics and placebo. We used novel visual and statistical methods to explore if smaller trials are associated with larger treatment effects and a selection model to explore if the probability of trial publication is associated with the magnitude of effect. We conducted a network meta-analysis of the published evidence as our primary analysis and used a sensitivity analysis considering low, moderate, and severe selection bias (that corresponds to the number of unpublished trials) with an aim to evaluate robustness of point estimates and ranking. We explored whether placebo-controlled and head-to-head trials are associated with different levels of PB.

Results: We found that small placebo-controlled trials exaggerated slightly the efficacy of antipsychotics, and PB was not unlikely in the evidence based on placebo-controlled trials; however, ranking of antipsychotics remained robust.

Conclusion: The total evidence comprises many head-to-head trials that do not appear to be prone to small-study effects or PB, and indirect evidence appears to ''wash out'' some of the biases in the placebo-controlled trials.

Keywords: Indirect comparison; Network meta-analysis; Publication bias; Selection model; Small study effect; Bayesian hierarchical model

1. Introduction

A well-documented example of publication bias (PB) in mental health is comparison of published and registered with the Food and Drug Administration (FDA) trials on the apparent efficacy of antidepressants by Turner et al. [1]. Turner et al. found great discrepancies in the results of published and unpublished but registered placebo-controlled trials. This prompted several reanalyses of existing trials in antidepressants and casted doubts about their efficacy as estimated in placebo-controlled trials [2–4]. As a result, Cipriani et al. [5] excluded data from placebo-controlled comparisons in the network meta-analysis (NMA) they conducted to evaluate the relative effectiveness of second-generation antidepressants. Turner et al. [6] also examined 24 placebo-controlled trials registered with the FDA involving eight FDA-approved antipsychotics for schizophrenia. They found an average 8% increase in the relative effectiveness of the antipsychotics in the published trials, although this increase was not found to be statistically significant.

In the decision-making context, we commonly encounter trials of many designs (sets of treatments compared in a study), and we aim to identify the most effective treatment. NMA allows a synthesis of trials with more than two treatments and provides an arsenal of methods and tools to visualize the evidence base and comprehend how information flows across multiple comparisons. It offers a series of advantages compared with simple meta-analysis such as increased power and precision, synthesis of direct and indirect evidence, evaluating relative effectiveness for pairs of

Funding: D.M. and G.S. received research funding from the European Research Council (IMMA 260559). O.E. received funding from Greek national funds through the Operational Program ''Education and Lifelong Learning'' of the National Strategic Reference Framework (NSRF)—Research Funding Program: ARISTEIA. Investing in knowledge society through the European Social Fund.

Conflict of interest: None.

* Corresponding author. Tel.: +30 26510 05694.
E-mail address: dmavridi@cc.uoi.gr (D. Mavridis).

http://dx.doi.org/10.1016/j.jclinepi.2015.05.027
0895-4356/

What is new?

Key findings

- Evidence from placebo-controlled trials comparing the least effective antipsychotics is distorted by small-study effects (SSEs) and publication bias (PB). This raises concerns about the validity of the results from network meta-analysis (NMA). We did not detect any SSEs or PB operating in head-to-head trials, and the compromised evidence based on placebo-controlled trials has immaterial impact on the estimated NMA effects for the most effective antipsychotics and their relative ranking. These findings enhance credence in NMA for appraising the overall efficacy and ranking of antipsychotics.

What this adds to what was known?

- This appraisal, using novel statistical methodology, exemplifies methods and ways of examining and adjusting NMA results for SSEs and PB. As there is often a dilemma in choosing between analyzing all trials or a subset that is believed to be less prone to bias, the methods used provide a compromise where all trials are used but the results are adjusted for SSEs or PB.

What is the implication and what should change now?

We suggest the following strategy when analyzing a network of interventions using NMA.

- Include all trials as a starting point (primary analysis).
- Apply statistical models to detect and, if necessary, adjust for PB or SSEs.
- In case of evidence of PB or SSEs, consider excluding certain comparisons from the analysis (eg, all placebo-controlled trials or very small/old trials) as a sensitivity analysis.

treatments for which there are no direct evidence and a ranking of interventions [7−10]. In light of the apparently distorted evidence from placebo-controlled trials for antidepressants and antipsychotics, we are concerned about the biases and validity in full NMA. However, little work has been done in exploring bias in head-to-head trials let alone in a network of interventions. Bias in certain comparisons (eg, placebo-controlled trials) may spread in the entire network and affect any relative effects of all treatment comparisons via indirect routes.

Systematic reviewers use a series of visual and statistical methods to explore PB [11−13]. A common strategy is to explore any differences in efficacy between small and large trials. Smaller trials are sometimes associated with larger efficacy, a phenomenon referred to as small-study effects (SSEs) [11−13]. Detecting SSE should not be directly equated with PB as there could be genuine reasons why smaller trials show larger treatment effects (eg, a treatment could be more beneficial to high-risk patients who are more difficult to recruit). Chaimani et al. [14] conducted a network meta-epidemiological study and found that small trials tended to exaggerate the effectiveness of the active interventions in placebo-controlled trials.

Selection of trials to be included in a review does not necessarily imply that an analysis of the included trials gives biased results. If published trials are a random sample of all trials conducted, then a meta-analysis will give unbiased results with increased uncertainty. The aim of this article was to evaluate the presence and possible impact of PB and SSE on the apparent efficacy of antipsychotic drugs derived from large meta-analyses that are likely to shape the decision-making process and explore how robust are the results to progressively more severe scenarios regarding the amount of unpublished evidence.

2. Methods

The data set is a network of 15 antipsychotic drugs and placebo for acute treatment in schizophrenia [15]. The primary outcome was efficacy as measured by mean overall change in symptoms in standardized scale and synthesized using standardized mean difference. A total of 167 trials were included, of which 130 were two-arm (43 placebo-controlled and 87 head-to-head trials), 35 were three-arm, and 2 were four-arm trials. The evidence network is presented in Appendix C/Fig. A.1 at www.jclinepi.com, and some of the characteristics associated with study size are presented in Table 1. The least effective drugs are compared mostly with placebo and not with other active agents, whereas the most effective treatments (with the exception of paliperidone) are compared mainly with haloperidol. Paliperidone and lurasidone are compared only in placebo-controlled two- or three-arm trials.

Understanding the flow of information in an NMA is essential when making judgments about the risk of bias that certain subgroups of trials might impose on the results. In an NMA, direct and indirect evidence are synthesized, and assumption or judgments about the risk of bias need to be derived considering the contribution of the various direct comparisons to the estimation of all relative treatment effects [16−19]. For instance, if placebo-controlled trials are subject to PB, the NMA treatment effects will be distorted for all comparisons, even for amisulpride, for which no placebo-controlled trials are available (Table 1). The last column in Table 1 presents the percentage contribution of placebo-controlled trials to various NMA estimates. (Details about the estimation