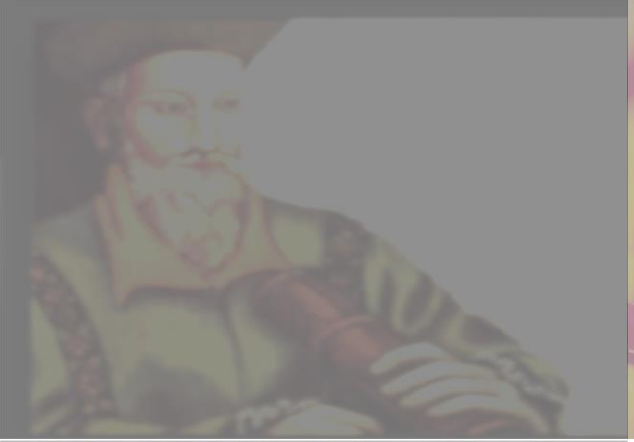


Publication bias in mental health clinical trials

Evidence Based Mental Health:
From Research to clinical practice
Corfu 2016

Dimitris Mavridis
University of Ioannina
Greece


Publication bias

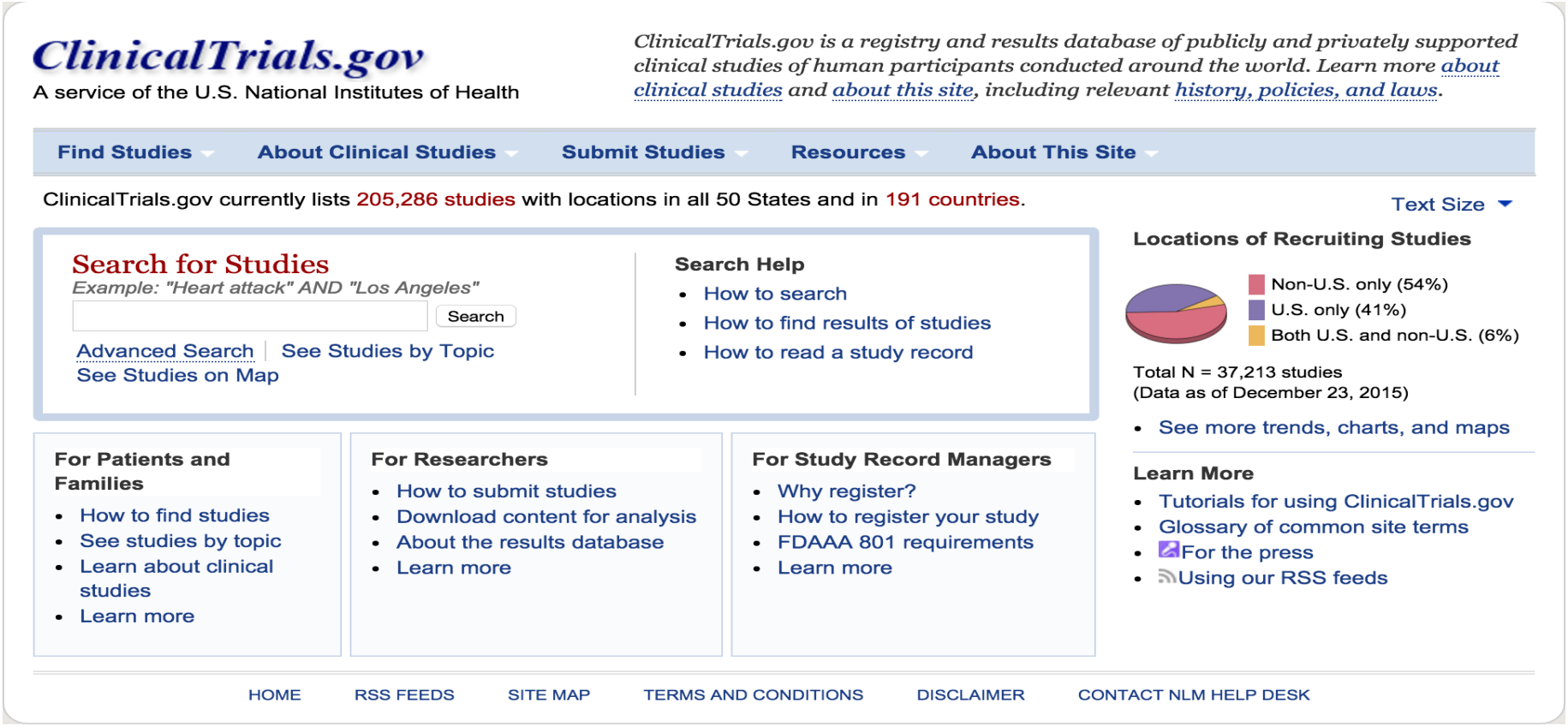
- Publication bias occurs when the publication of research results depends not just on the quality of the research but on its nature and direction (Dickersin 1990)
- There is a tendency in medical journals to publish more frequently studies showing significant results to studies not suggesting significant interaction effects
- It has been found that the most common reason for non-publication is an investigators declining to submit results for publication (Easterbrook et al 1991)
- The evidence published represent a biased/non-representative sample of the overall evidence
- Publication bias will result in exaggerated intervention effects

Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. JAMA 263(10):1385-1389
Easterbrook, P. J.; Berlin, J. A.; Gopalan, R.; Matthews, D. R. (1991). Publication bias in clinical research. Lancet 337 (8746)

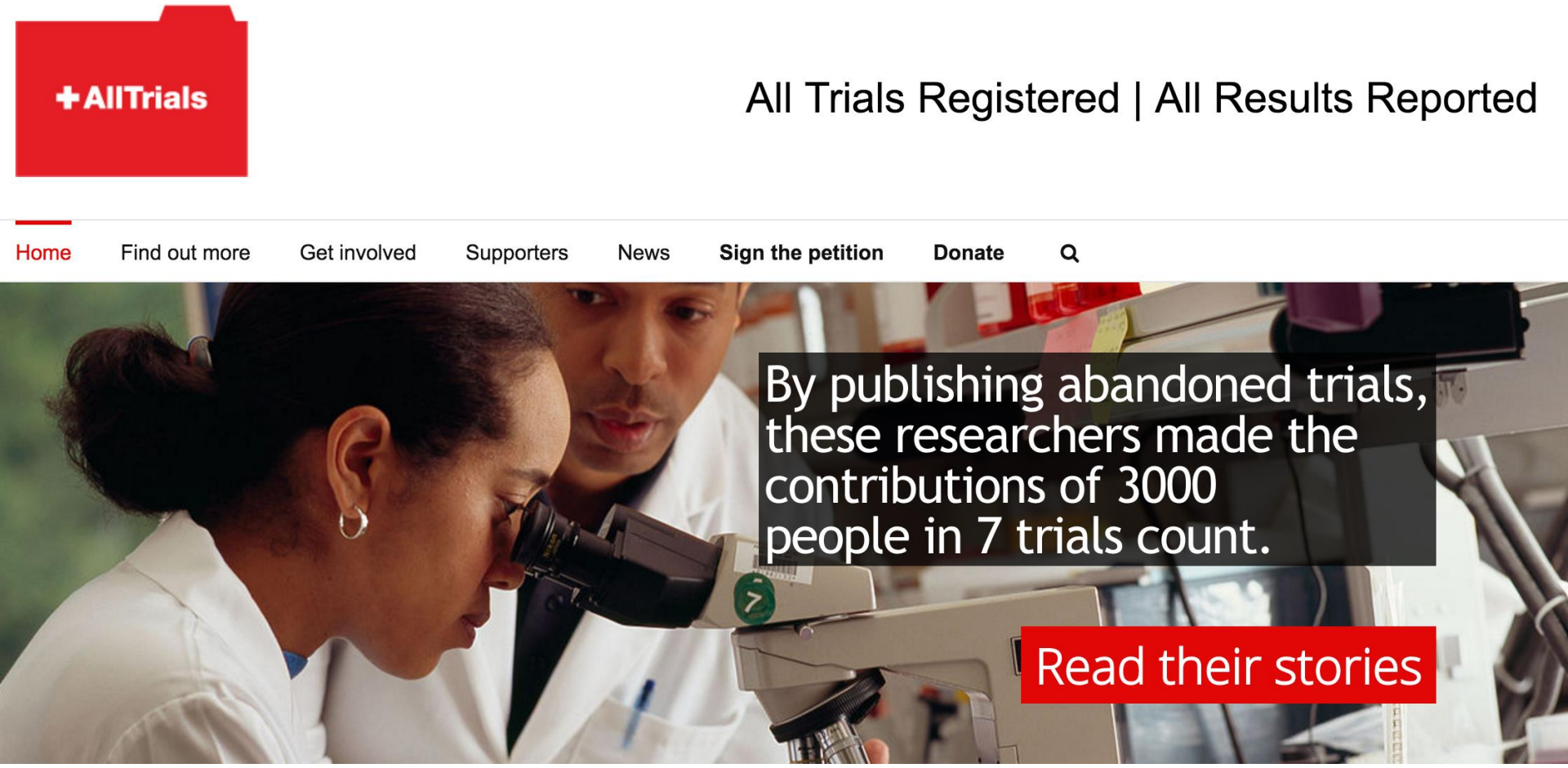
Avoiding publication bias

- Investigate 'grey literature' (reports by government, academics, industry and individuals not published in scientific journals)
- Pre-emptive strategies
- Trial registries : Major medical journals require researchers register their trials in public trials registry

Trial registries



alltrials.net



Trial registries

- Prospective registration of studies and public disclosure of their results
- less than a third of medical journals require or encourage trial registration (Wager et al 2013)
- at least 50% of registered trials did not report results within a year of completion (Prayle et al 2012)
- Most drugs were included in trials conducted when registration was not mandatory (Goldacre 2013)

Goldacre B. Are clinical trial data shared sufficiently today? No 2013; 347
Prayle AP, Hurley MN, Smyth AR. Compliance with mandatory reporting of clinical trial results on ClinicalTrials.gov: cross sectional study. BMJ. 2012;344:d7373.
Wager, E., Williams, P. (2013). Hardly worth the effort? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. BMJ 34

Small study effects

- Small studies are less precise and are expected to be more widely scattered around the mean
- Small study effects happens when smaller studies have systematically larger effects compared to the large studies
- Small study effects should not necessarily be equated with publication bias although it is usually a good proxy for it

Egger M, Smith GD, Scheider M, Minder C. Bias in meta analysis detected by a simple graphical test. *British Medical Journal.* 1997;**315**:629-34.
Mavridis D, Salanti G. Exploring and accounting for publication bias in mental health: a brief overview of methods. *Evidence Based Mental Health* 2014; **17**(1):11-15
Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;**323**(7304):101-5.

Small study effects due to heterogeneity

- Smaller studies tend to be conducted with less mathematical and methodological rigor
- Patients at high risk may benefit substantially from the treatment and they are more likely to be included in a small study because they are harder to recruit
- Trials in populations where the intervention is very effective require smaller sample sizes to achieve pre-specified levels of power

Possible reasons for small-study effects

Adapted from Egger et al. (Egger 1997a).

1. **Selection biases**:
   - Publication bias:
     - Delayed publication (also known as 'time-lag' or 'pipeline') bias.
     - Location biases:
       - Language bias;
       - Citation bias;
       - Multiple publication bias.
   - Selective outcome reporting.
2. **Poor methodological quality leading to spuriously inflated effects in smaller studies**:
   - Poor methodological design;
   - Inadequate analysis;
   - Fraud.
3. **True heterogeneity**:
   - Size of effect differs according to study size (for example, due to differences in the intensity of interventions or differences in underlying risk between studies of different sizes).
4. **Artefactual**:
   - In some circumstances (see Section 10.4.3), sampling variation can lead to an association between the intervention effect and its standard error.
5. **Chance**.

Suggested remedies for identifying and accounting for Publication Bias

- A plethora of naïve methods (fail-safe method, excess statistical significance, trim and fill) – not recommended
- Visual (funnel plot) methods (aim at small-study effects)
- Regression-based methods (aim at small-study effects)
- Use a model to describe the study selection process (aim at publication bias)

Turner's work on antidepressants

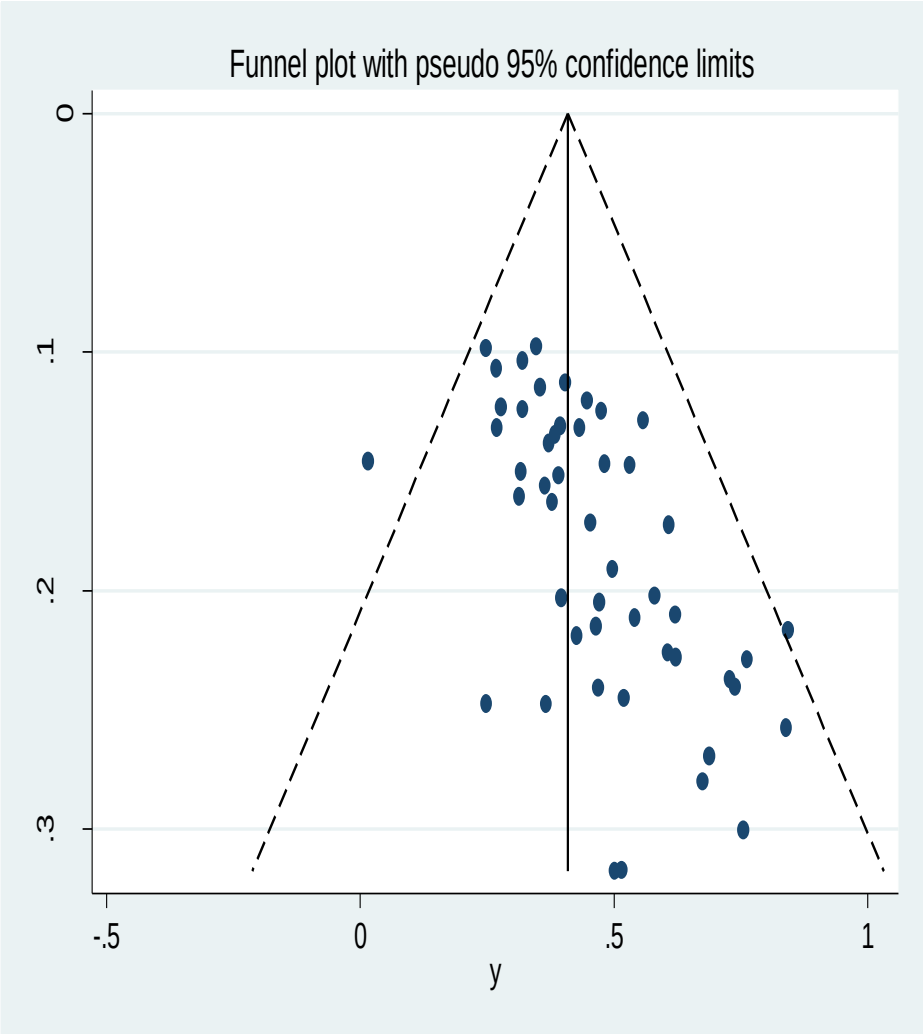
- Turner et al found 73 studies registered with the FDA used for the licensing of antidepressants drugs between 1987 and 2004 involving 12 drugs.
- 50 studies of these 73 studies were subsequently published in medical journals
- From the 38 FDA studies with statistically significant results only one was not published
- from the 36 FDA with non-statistically significant results only three were published
- another 11 were published with results conflicting those presented in the FDA report

Turner EH, Mathews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine.* 2008, **358**(3):252-260.

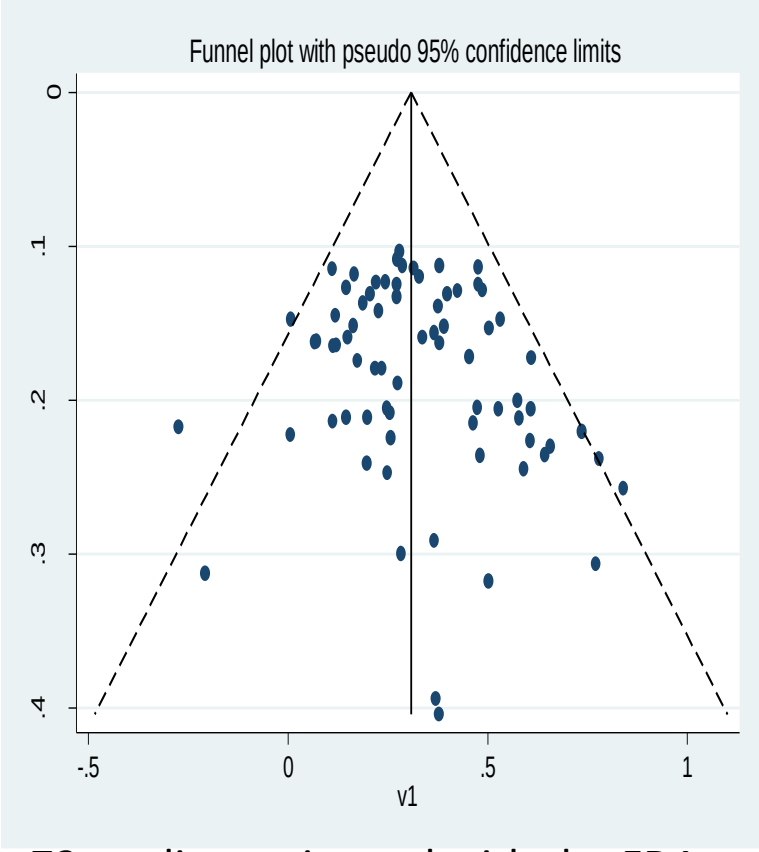
Funnel plot

- A scatter plot of the intervention effect estimates from individual studies against some measure of each study's precision

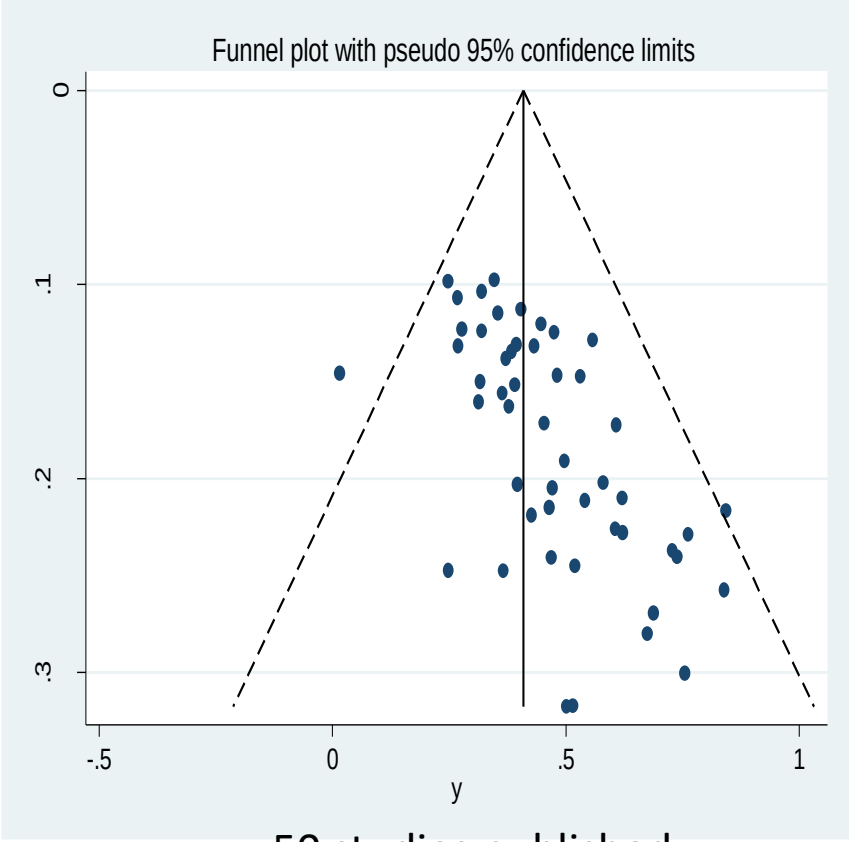
- We check for asymmetries in the funnel plot (a gap in the bottom corner)



Funnel plots without and with publication bias



73 studies registered with the FDA



50 studies published

Funnel plot asymmetry

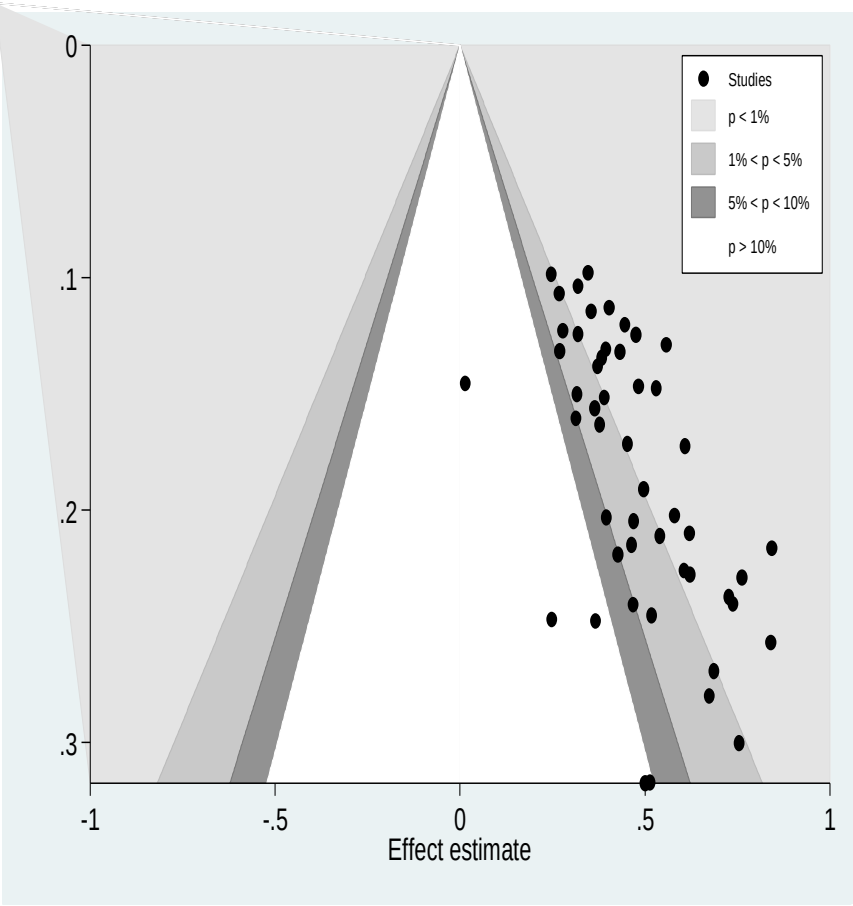
- An indication of small-study effects
- may be due

➢ Publication bias (small studies suggesting a not significant interaction effect do not get published)

➢ heterogeneity

Contour-enhanced funnel plots

- Includes contour lines corresponding to different levels of statistical significance
- If non-significant studies are missing this is probably due to publication bias



Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. Journal of Clinical Epidemiology 2008; 61(10):991-996.

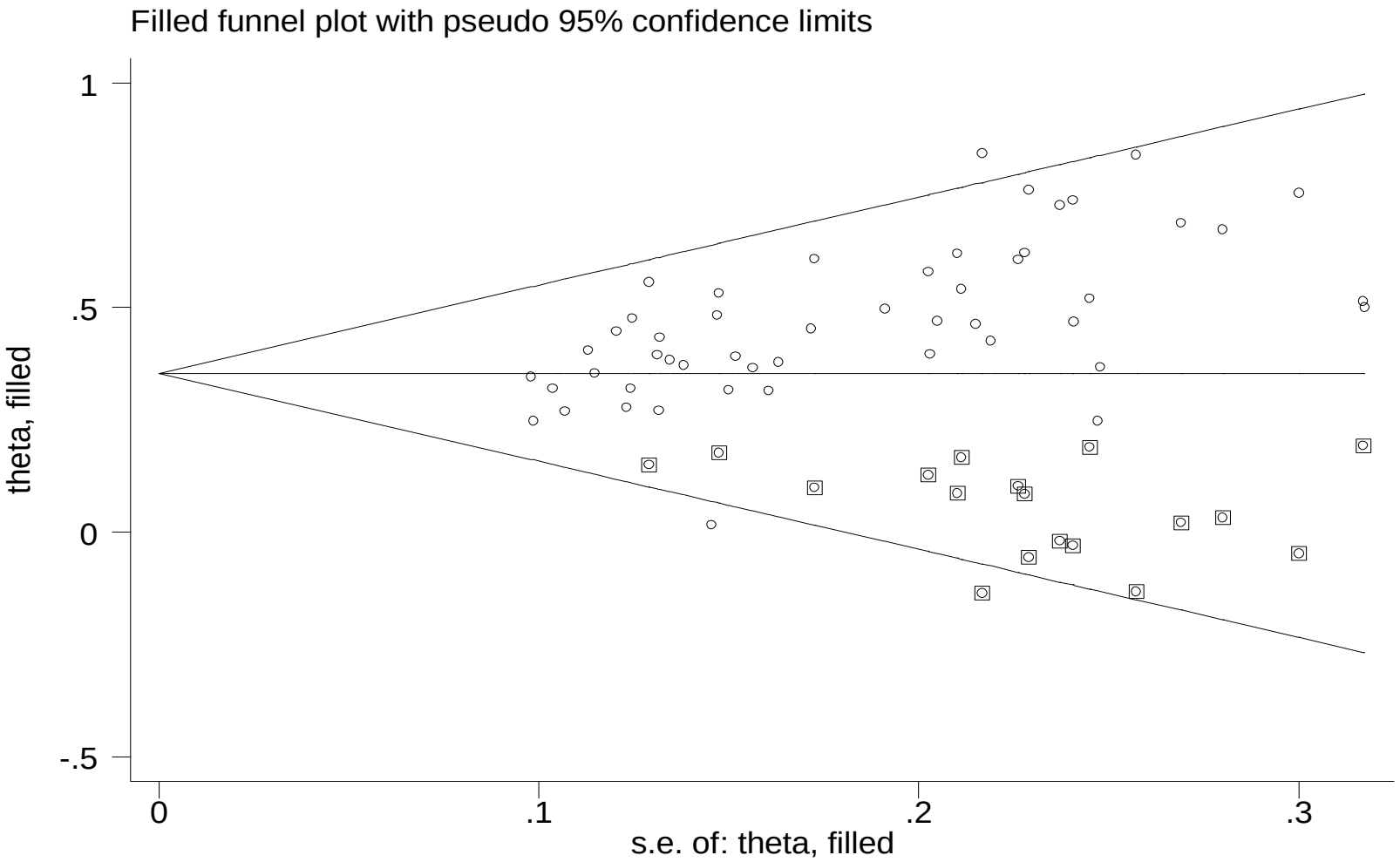
Fail-safe N method (Not recommended)

- Suppose we get a significant p-value based on a meta-analysis with k studies
- How many studies we would need to incorporate in the analysis before the p-value becomes non-significant
- If a small number is needed (say 5 or 10) then the true effect may be non-significant
- If a large number is needed (say 20000) then the true effect is most probably significant

Trim and fill (Not recommended)

- Trim-of the asymmetric right-hand side of a funnel plot after estimating the number of studies in this group.
- Use the symmetric remainder to estimate the true center
- Replace trimmed studies and their missing counterparts around the center.
- Estimate the effect size and its variance based on the new funnel plot.

Funnel plot with imputed studies



Limitations of the trim and fill method

- Assumes there is a symmetric funnel plot
- Assumes the only reason for funnel plot asymmetry is publication bias
- Imputes studies as if they were observed
- Behaves poorly in the presence of heterogeneity

Regression-based methods – tests for funnel plot asymmetry

- Statistical analogues for testing funnel plot asymmetry
- Assess the degree of association between the study effect and its precision
- This association can be illustrated by drawing a regression line in the funnel plot

- H_0: there is no association between effect size and standard error

Regression-based methods

N_{tot} is the total sample size, N_E and N_C are the sizes of the experimental and control intervention groups, S is the total number of events across both groups and $F = N_{tot} - S$. Note that only the first three of these tests (Begg 1994, Egger 1997a, Tang 2000) can be used for continuous outcomes.

Reference	Basis of test
(Begg 1994)	Rank correlation between standardized intervention effect and its standard error.
(Egger 1997a)	Linear regression of intervention effect estimate against its standard error, weighted by the inverse of the variance of the intervention effect estimate.
(Tang 2000)	Linear regression of intervention effect estimate on $1/\sqrt{N_{tot}}$, with weights N_{tot}.
(Macaskill 2001)*	Linear regression of intervention effect estimate on N_{tot}, with weights $S \times F/N_{tot}$.
(Deeks 2005)*	Linear regression of log odds ratio on $1/\sqrt{ESS}$ with weights ESS, where effective sample size $ESS = 4N_E \times N_C/N_{tot}$.
(Harbord 2006)*	Modified version of the test proposed by Egger et al., based on the 'score' (O–E) and 'score variance' (V) of the log odds ratio.
(Peters 2006)*	Linear regression of intervention effect estimate on $1/N_{tot}$, with weights $S \times F/N_{tot}$.
(Schwarzer 2007)*	Rank correlation test, using mean and variance of the non-central hypergeometric distribution.
(Rücker 2008)	Test based on arcsine transformation of observed risks, with explicit modelling of between-study heterogeneity.

Egger's regression test

- Detects funnel plot asymmetry
- A weighted linear regression with standard error (se) as covariate

$$y_i = a + bse_i + e_i$$

weighted by $w_i = \frac{1}{se_i^2}$

- Significant values for β indicate different effects for smaller studies

Recommendations

- Ideally, there should be at least 10 studies to apply a regression-based method. Otherwise, the tests have low power in detecting a real asymmetry from chance
- Tests are conservative, a non-significant result does not necessarily mean absence of publication bias

A missing data problem

- Publication bias is a missing data problem
- If the reason studies are missing is related to the outcome of the studies, data are missing not at random (MNAR)
- Any analysis on observed data when missing data are MNAR would give biased results
- We can only resort to assumptions about missing data and conduct a sensitivity analysis

Selection models

- There are two models considered and combined

➢ Measurement model : Specifies the distribution of the effect size when there is no publication bias

➢ Selection model : Specifies the mechanism by which effect sizes are selected to be observed

Sutton AJ, Song F, Gilbody SM, Abrams KR: Moddeling publication bias in meta-analysis : a review. *Statistical Methods in Medical Research* 2000; **9:** 421-445.

Selection models

- Copas suggested a selection model where the probability of publication of a study depends on its standard error

- He assumes

➢ the larger the sample size the larger the probability of publication

➢ The larger the effect size the larger the probability of publication

Selection models

- Copas suggested a selection model where the probability of publication of a study depends on its standard error

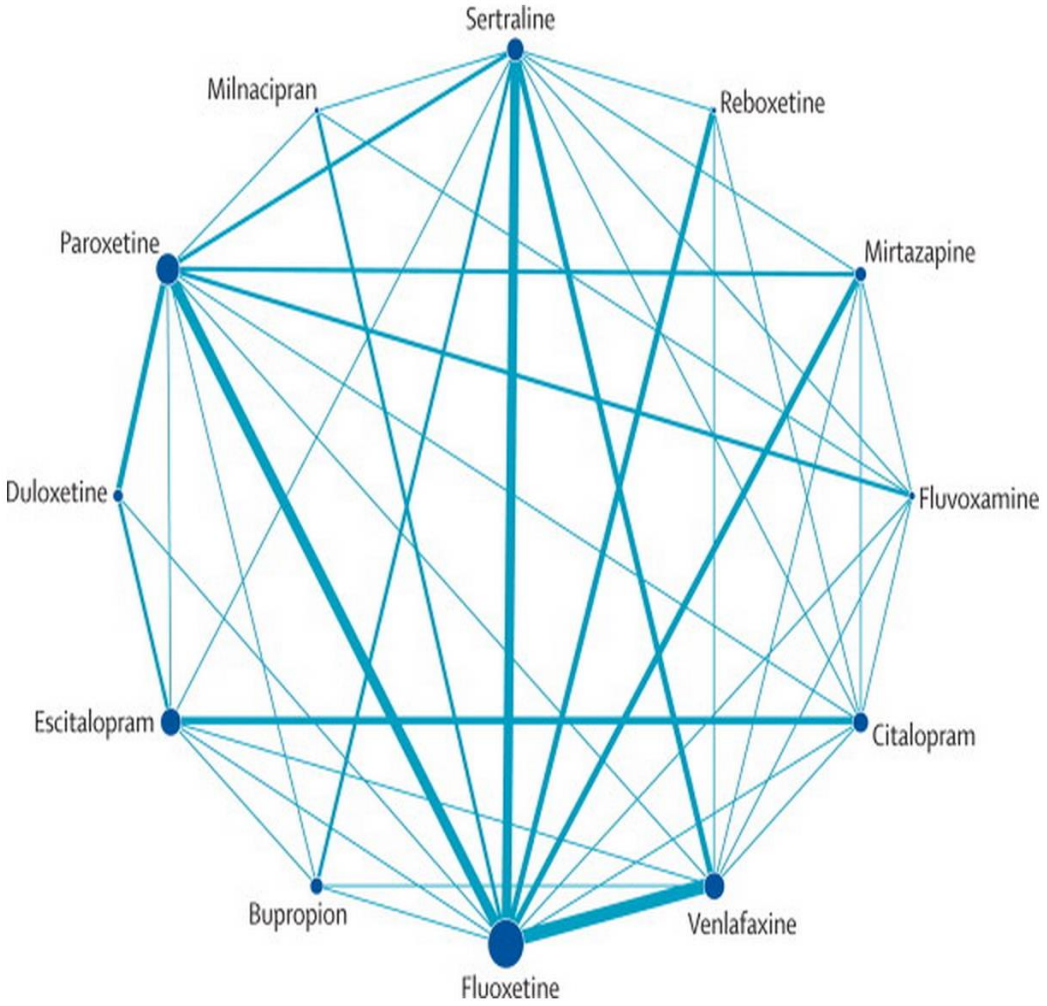
- He assumes
  - the larger the sample size the larger the probability of publication
  - The larger the effect size the larger the probability of publication

Copas J, Shi JQ: Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics* 2000; **1**(3): 247-262.
Mavridis D, Welton NJ, Sutton A and Salanti G. A selection model for accounting for publication bias in a full network meta-analysis. *Statistics in Medicine*. 2014
Mavridis D, Salanti G. Exploring and accounting for publication bias in mental health: a brief overview of methods. *Evidence Based Mental Health* 2014; **17**(1):11-15.

Network plot of depression trials

Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet.* 2009;**373**(9665):746-58.

- Network meta-analysis synthesizes both direct and indirect evidence
- Excluded placebo-control trials because they show exaggerated results
- A biased treatment comparison will contaminate other effect estimates through indirect evidence
- 117 RCT's comparing active antidepressants
- Certain antidepressants are more effective and better tolerated than others



Turner's work on antipsychotics

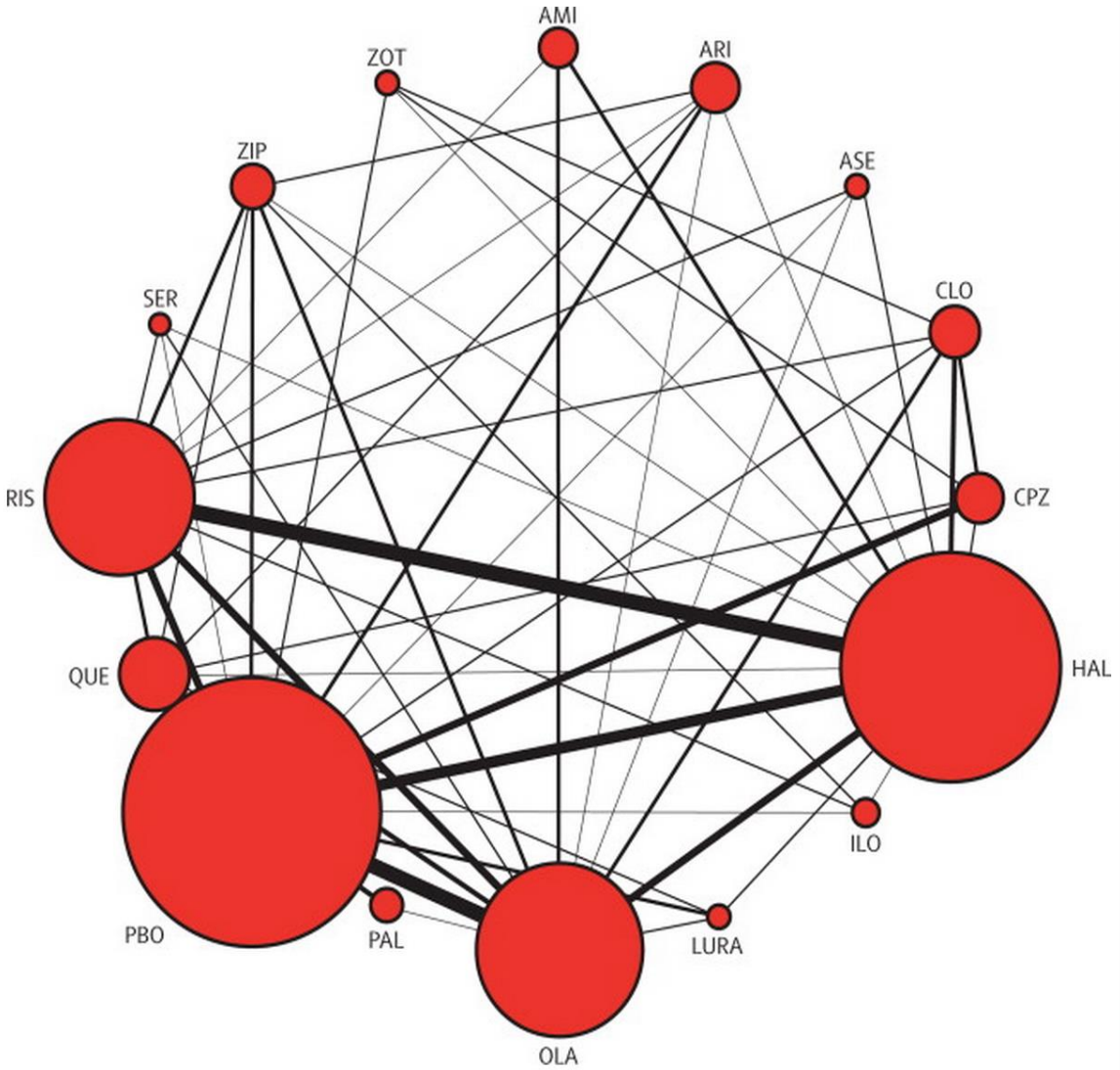
- Identify 24 trials registered with the FDA comparing 8 second-generation antipsychotics to placebo
- Only 4 trials remain unpublished
- All antipsychotics were considered equal
- The increase in the summary estimate in the published trials was modest and not significant (8%)
- The four unpublished trials had an effect size (0.23,95% CI 0.07,0.39) less than half of that of the published trials (0.47,95% CI 0.40,0.54)
- Publication bias is not severe possibly because antipsychotics demonstrate superiority to placebo more consistently

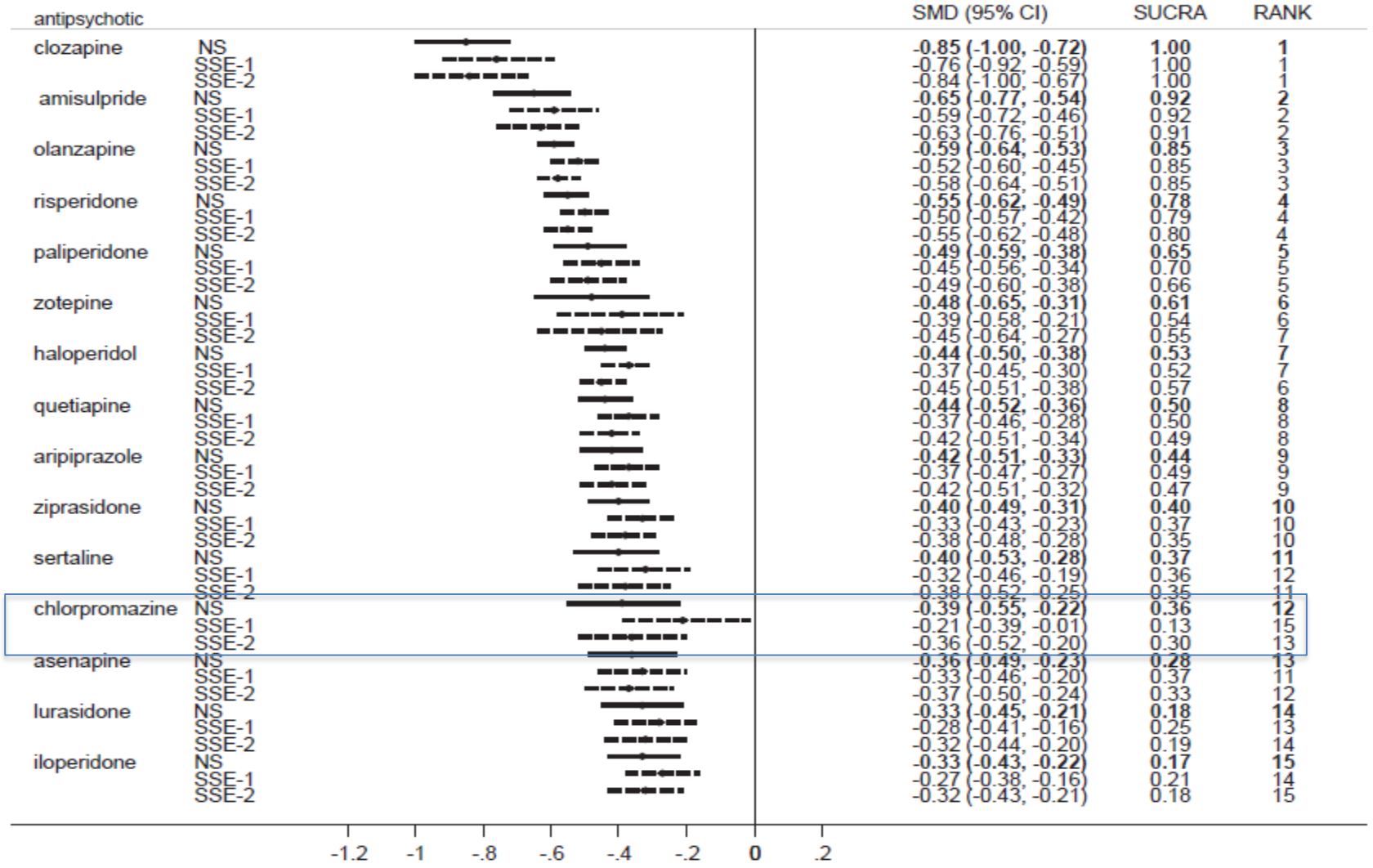
Turner EH, Knoepflmacher D, Shapley L. Publication bias in antipsychotic trials: an analysis of efficacy comparing the published literature to the US Food and Drug Administration database. *PLoS Medicine.* 2012;**9**(3):e1001189.

Network plot of antipsychotic trials

Leucht S, Cipriani A, Spineli L, Mavridis D, et al. Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet.* 2013;**382**(9896):951-962.

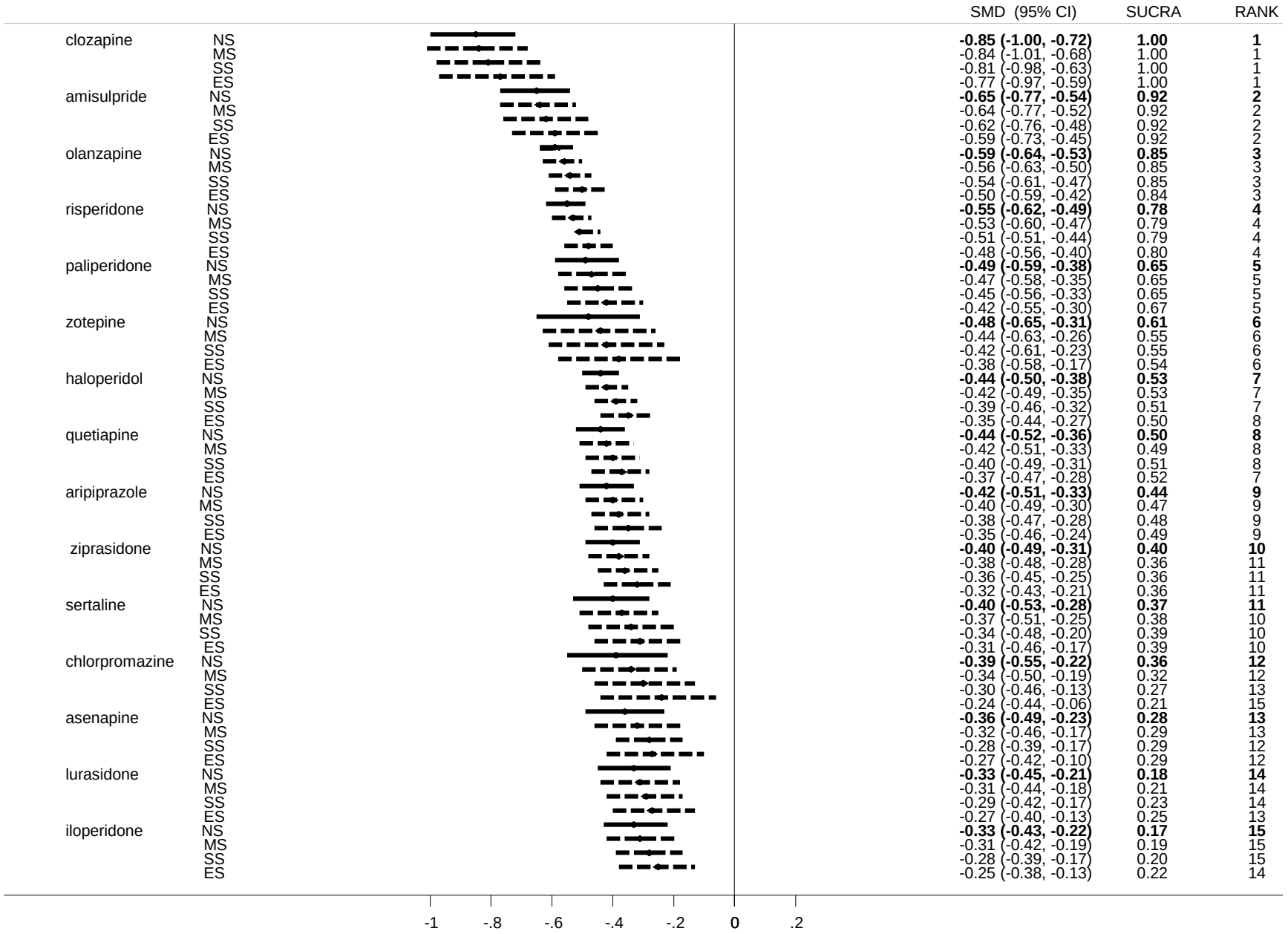
- The least effective drugs are mainly concerned to placebo
- 73 placebo-control trials and 92 head-to-head trials
- PB and SSE were detected in placebo control trials but not in head-to-head trials





Mavridis, D., Efthimiou, O., Leucht, S. and Salanti, G. Publication bias and small-study effects magnified effectiveness of antipsychotics but their relative efficacy remained invariant. Journal of Clinical Epidemiology 2015

Summary estimates with 95% C.I. (compared to placebo), SUCRA values and rank.
NS: no selection, MS=moderate selection, SS: severe selection, ES: extreme selection



Correlations between magnitude of effect and propensity for publication

Table 2. Correlations between the probability a trial is included in a meta-analysis and the magnitude of effect it provides (standardized mean difference), and heterogeneity standard deviation estimates for the four publication scenarios

Comparison	No selection bias	Moderate selection bias	Severe selection bias	Extreme selection bias
Placebo controlled	0.03 (−0.93, 0.96)	−0.41 (−0.97, 0.41)	−0.51 (−0.92, 0.08)	**−0.43 (−0.75, −0.06)**
Head to head	0.00 (−0.93, 0.95)	0.18 (−0.41, 0.68)	0.15 (−0.23, 0.51)	0.11 (−0.13, 0.33)
Multiarm: placebo-controlled comparison	−0.06 (−0.52, 0.40)	−0.21 (−0.57, 0.15)	**−0.31 (−0.59, −0.03)**	**−0.33 (−0.53, −0.13)**
Multiarm: head-to-head comparison	−0.02 (−0.18, 0.13)	−0.03 (−0.18, 0.13)	−0.02 (−0.18, 0.14)	−0.02 (−0.19, 0.14)
Heterogeneity standard deviation	0.09 (0.04, 0.13)	0.10 (0.05, 0.15)	0.11 (0.06, 0.15)	0.10 (0.04, 0.15)

Different correlation coefficients are assumed for placebo-controlled (43 studies), head-to-head (87 studies), and 35 three-arm trials. 95% Credible intervals are given in parentheses. Significant results are given in bold.

Conclusions regarding the analysis of psychotic trials

- Results from a NMA of 167 psychotic trials seem to be robust to the amount of unpublished evidence considered
- Efficacy was not reduced significantly and ranking was robust especially for the most effective treatments
- Excluding placebo-control trials would give similar estimates but with increased uncertainty
- The large number of head-to-head trials washes out some of the bias caused by small placebo-control trials
- Publication bias in psychotic trials is less severe compared to antidepressant trials probably because of the superiority of active drugs compared to placebo

Key findings

- Include all trials as a starting point (primary analysis)
- If there is a considerable number of trials (e.g. 10)
- Use visual methods (funnel plots)
- Apply statistical models (selection and regression-based models) to explore how robust effect estimates are

Evid Based Mental Health 2014;**17**:11-15 doi:10.1136/eb-2013-101700

Statistics in practice

Exploring and accounting for publication bias in mental health: a brief overview of methods

Editor's choice

Dimitris Mavridis[1,2], **Georgia Salanti**[1]

+ Author Affiliations

Correspondence to

Dr Dimitris Mavridis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; dimi.mavridis@googlemail.com

Abstract

Objective Publication bias undermines the integrity of published research. The aim of this paper is to present a synopsis of methods for exploring and accounting for publication bias.

Methods We discussed the main features of the following methods to assess publication bias: funnel plot analysis; trim-and-fill methods; regression techniques and selection models. We applied these methods to a well-known example of antidepressants trials that compared trials submitted to the Food and Drug Administration (FDA) for regulatory approval.

Results The funnel plot-related methods (visual inspection, trim-and-fill, regression models) revealed an association between effect size and SE. Contours of statistical significance showed that asymmetry in the funnel plot is probably due to publication bias. Selection model found a significant correlation between effect size and propensity for publication.

Conclusions Researchers should always consider the possible impact of publication bias. Funnel plot-related methods should be seen as a means of examining for small-study effects and not be directly equated with publication bias. Possible causes for funnel plot asymmetry should be explored. Contours of statistical significance may help disentangle whether asymmetry in a funnel plot is caused by publication bias or not. Selection models, although underused, could be useful resource when publication bias and heterogeneity are suspected because they address directly the problem of publication bias and not that of small-study effects.

This Article

- Full text
- PDF

- Services
- Email this link to a friend
- Alert me when this article is cited
- Alert me if a correction is posted
- Alert me when eletters are published
- Article Usage Statistics
- Similar articles in this journal
- Similar articles in PubMed
- Add article to my folders
- Download to citation manager
- Request permissions
- Add to portfolio

- Responses
- Submit a response
- No responses published

+ Citing articles

+ Google Scholar

+ PubMed

- Related Content
- EBMH Statistics in Practice
- Editor's choice

+ Social bookmarking

Online first | Current | Most read

- Outcomes:Three simple questions

Evidence-Based Mental Health

Improving clinical practice for psychiatrists and psychologists

Evid Based Mental Health 2014;**17**:30 doi:10.1136/eb-2013-101699

Statistics in pills

Statistics in pills

How to assess publication bias: funnel plot, trim-and-fill method and selection models

doi:10.1136/eb-2013-101699

This new section of the Journal is aimed at providing the essential information readers should know about the topics that are addressed in the "Statistics in practice" paper published in the same issue of the journal. This stand-alone section has to be seen as an articulated summary of the main notions clinicians have to know about some basic concepts in statistics, which may be useful for their evidence based practice. After going through these notes, readers are encouraged to read the "Statistics in practice" articles. Of course, we welcome any feedback from you (via email or Twitter) about this!
The EBMH Editors

FUNNEL PLOT

A funnel plot is a scatter plot of the treatment effect estimates from individual trials against a measure of study's precision (usually the standard error (SE)).[1]

- Asymmetry in the funnel plot should not be automatically equated with publications bias.
- There are various alternative explanations for funnel plot asymmetry, such as heterogeneity, selective outcome reporting and chance.
- A sizeable number of studies (by convention, usually at least 10 though more may be needed in the presence of substantial heterogeneity) are needed to obtain a visual assessment of the funnel plot.
- Funnel plots should be better seen as a means for exploring small-study effects; small studies showing systematically larger effects than large studies.
- Adding contours of statistical significance to the funnel plot may help distinguish publication bias from other causes of funnel plot asymmetry.

TRIM-AND-FILL METHOD

The trim-and-fill is a funnel plot-derived, two-step method aimed at both identifying publication bias and adjusting results for it.[1] Phase 1 (Trimming): to exclude small studies in order to have a symmetrical plot and then estimate an adjusted summary effect considering only the larger studies. Phase 2 (Filling): to replicate the funnel plot replacing the excluded studies with their 'missing' counterparts around the adjusted summary estimate.

- The trim-and-fill method provides a summary effect adjusted for publication bias.
- It allows estimating the number of unpublished studies.
- The trim-and-fill method assumes publication bias as the only reason for funnel plot asymmetry which is an unrealistic assumption. It should be used as a sensitivity analysis as its inventors suggested.

SELECTION MODELS

Selection models focus on the selection process, that is, the mechanism by which trials are selected for publication.[1] Using selection models, researchers can estimate the likely impact the missing studies would have, had they been included in the meta-analysis. One of the key assumptions in the selection models is that the included sample of studies is not at random. The studies have been included because they have some characteristics that increase their propensity for publication, therefore the overall estimate is conditional to the observed studies that have been published and identified. Taking this into account, it is possible to calculate the marginal effect size, which is the effect size unconditional to the publication status.

- Assumptions are needed about the factors that influence the probability of publication for a study. The probability of publication is typically assumed to be a function of a study's p value or sample size.
- A selection process does not necessarily entail bias. If the effect sizes are comparable in small and large studies, then even a strong selection process will not alter the results of meta-analysis.
- In meta-analysis, selection models condition the observed effect size in each study to its propensity for publication and they estimate the unconditional summary that pertains to all studies that have been carried out, either published or unpublished.
- Copas selection model is the most sophisticated selection model because it assumes that probability of publication depends on both the study's effect size and SE. It allows us to estimate the correlation between the probability of publication of a study and the effect size. If different from zero, this means that the selection process has produced a publication bias.
- Unlike approaches based on funnel plot asymmetry, selection models test and adjust for publication bias, without being confounded by heterogeneity.
- The mechanism of the selection process is unknown and a sensitivity analysis is advocated in which the intervention effect is estimated under different assumptions about the severity of selection bias.

Dimitris Mavridis,[1,2] Georgia Salanti[1]

[1]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece
[2]Department of Primary Education, University of Ioannina, Ioannina, Greece

Correspondence to: Dr Dimitris Mavridis, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece; dimi.mavridis@googlemail.com

Funding Dimitris Mavridis and Georgia Salanti received research funding from the European Research Council (IMMA 260559).

Competing interests None.

REFERENCE

1. **Mavridis D,** Salanti G. Exploring and accounting for publication bias in mental health: a brief overview of methods. *EBMH* 2014;**1**.