University of BRISTOL

Dealing with small studies

Julian Higgins

School of Social and Community Medicine, University of Bristol, UK

University of BRISTOL

Outline

- Theoretical problems with small studies: estimation of within-study variance
- Some empirical evidence relating to small studies
- Empirical problems of small studies: small study effects
  - exclusion of small studies
  - forest plots (including cumulative versions)
  - funnel plots (including contour-enhanced versions)
  - regression approaches
  - alternative weighting approaches
- Some remarks on small *meta-analyses*

University of BRISTOL

Theoretical problems with small studies: estimation of within-study variance

Acknowledgements: Mark Simmonds

Models for binary data (including logistic regression)

- All the methods in RevMan are two-stage methods
  - obtain an estimate of treatment effect from each study
  - compute a weighted average of these estimates

- An alternative is a one-stage method
  - usually this requires individual participant data
  - but we can create this for dichotomous outcomes

University of BRISTOL

IPD from a 2×2 table

Study 1 (n_1 people)

	Alive	Dead
Treatment	rt_1	ft_1
Control	rc_1	fc_1

Person	Study	Group	Dead	
1	1	0	0	
2	1	0	⋮	rt_1
⋮	1	⋮	0	
⋮	1	⋮	1	
⋮	1	⋮	⋮	ft_1
⋮	1	0	1	
⋮	1	1	0	
⋮	1	1	⋮	rc_1
⋮	1	⋮	0	
⋮	1	⋮	1	
⋮	1	⋮	⋮	fc_1
n_1	1	1	1	

University of BRISTOL

Logistic regression

- With individual participant data, we can use the methods we use to analyse a single study
  - except that we don't have participant-level characteristics

- In particular we can use logistic regression
  - but we can't adjust for things like age, sex, severity

- A fixed-effect meta-analysis can be done using logistic regression, stratifying by study (a dummy variable for each study)
- A random-effects meta-analysis can be done using random-effects logistic regression (**meqrlogit**, previously xtmelogit)
  - see Simmonds and Higgins, *Statistical Methods in Medical Research* (early view online)

University of BRISTOL

Stata: meta-analysis using random-effects logistic regression

- With data in variables *study*, *rt* ,*nt*, *rc*, *nc*

```
reshape long n r, i(author) j(trt) string
gen treat = 0
replace treat=1 if trt=="t"
xi: meqrlogit s i.study i.treat || study: treat, nocons binomial(n)
```

- For out example (haloperidol) variables are *author*, *rh*, *fh*, *rp*, *fp*

```
gen nh = rh+fh
gen np = rp+fp
reshape long n r, i(author) j(trt) string
gen treat = 0
replace treat=1 if trt=="h"
xi: meqrlogit r i.author i.treat || author: treat, nocons binomial(n)
```

University of BRISTOL

Comparison of methods

Method for tau	Method for confidence interval	Meta-analysis result (95% CI)	Estimate of τ^2
Fixed-effect analysis		OR = 2.85 (1.99, 4.10)	n/a
DerSimonian-Laird	Z (normal)	OR = 4.20 (2.42, 7.30)	0.48
DerSimonian-Laird	Hartung-Knapp (t)	OR = 4.20 (2.31, 7.64)	"
Paule-Mandel (empirical Bayes)	Z (normal)	OR = 4.14 (2.40, 7.13)	0.45
Paule-Mandel (empirical Bayes)	Hartung-Knapp (t)	OR = 4.14 (2.30, 7.45)	"
REML	Z (normal)	OR = 4.38 (2.44, 7.86)	0.60
REML	Hartung-Knapp (t)	OR = 4.38 (2.33, 8.24)	
Profile likelihood	Z (normal)	OR = 4.25 (2.39, 8.82)	0.51
Random-effects logistic regression		**OR = 4.72 (2.61, 8.53)**	**0.25**

University of BRISTOL

What about continuous data?

- For the individual study results, we should be using t-distributions for the confidence intervals

- But I don't know of discussions of adjustment to weights in meta-analysis to account for small study issues

University of BRISTOL

Empirical problems of small studies: small study effects

BMJ

BMJ 2013;346:f2304 doi: 10.1136/bmj.f2304 (Published 24 April 2013)

RESEARCH

Influence of trial sample size on treatment effect estimates: meta-epidemiological study

Agnes
Isabell
directo

[1]INSERM,
Paris, Fra
University

Abstrac

Objective
effect esti

Design M

Data sou
assessing
the 10 lea
Citation R

Data extr
from each

Data synt
sample si
with 25%
and using
patients).
between
ratios (wh
in smaller

Results T
trials, rega
the larges
trials in qu
0.68, 95%
2 (0.83, 0.
0.95). Sim
estimates
patients or more, treatment effects were, on average, 48% larger in trials

Comparison	No (%) of meta-analyses included in analysis (n=93)	No (%) of randomised controlled trials included in analysis (n=735)	Ratio of odds ratios (95% CI)	τ^2
Quarters (Q) within meta-analyses				
Q1 *v* Q2-4	88 (95%)*	711 (97%)	0.77 (0.65 to 0.91)	0.11
Q1 and Q2 *v* Q3 and Q4	93 (100%)	735 (100%)	0.81 (0.74 to 0.88)	0.02
Q1-3 *v* Q4	92 (99%)†	722 (98%)	0.85 (0.79 to 0.90)	0.00
Fixed thresholds across meta-analyses (based on no of patients)				
<50 *v* ≥50	52 (56%)	417 (57%)	0.56 (0.45 to 0.70)	0.00
<100 *v* ≥100	72 (77%)	602 (82%)	0.74 (0.65 to 0.85)	0.00
<200 *v* ≥200	67 (72%)	598 (81%)	0.76 (0.68 to 0.85)	0.02
<500 *v* ≥500	43 (46%)	434 (59%)	0.81 (0.74 to 0.89)	0.00
<1000 *v* ≥1000	28 (30%)	301 (41%)	0.82 (0.76 to 0.90)	0.00

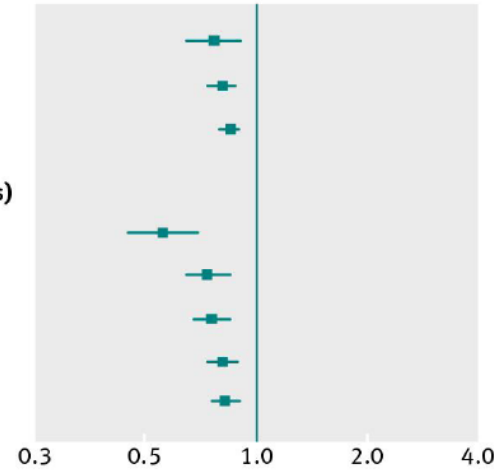


Fig 2 Comparison of treatment effect estimates between trial sample sizes grouped by quarters (from quarter 1 with the smallest trials, to quarter 4 with the largest trials) and by fixed thresholds (according to numbers of patients). All 93 meta-analyses did not contribute to the analysis, depending on the threshold used (at least one trial with a sample size less and more than the threshold within each meta-analysis was required to perform these analyses). *Five meta-analyses did not contribute to the analysis (three included fewer than four trials; for two trials, the meta-regression model did not converge, therefore, the ratio of odds ratios could not be obtained). †One meta-analysis did not contribute to the analysis (the meta-regression model did not converge, therefore, the ratio of odds ratios could not be obtained)

"Treatment effect estimates were significantly larger in smaller trials, regardless of sample size."

OPEN ACCESS Freely available online

PLOS ONE

The Impact of Study Size on Meta-analyses: Examination of Underpowered Studies in Cochrane Reviews

Rebecca M. Turner[1]*, Sheila M. Bird[1], Julian P. T. Higgins[2,3]

1 MRC Biostatistics Unit, Institute of Public Health, Cambridge, United Kingdom, **2** School of Social and Community Medicine, University of Bristol, Bristol, United Kingdom, **3** Centre for Reviews and Dissemination, University of York, York, United Kingdom

Abstract

Background: Most meta-analyses include data from one or more small studies that, individually, do not have power to detect an intervention effect. The relative influence of adequately powered and underpowered studies in published meta-analyses has not previously been explored. We examine the distribution of power available in studies within meta-analyses published in Cochrane reviews, and investigate the impact of underpowered studies on meta-analysis results.

Methods and Findings: For 14,886 meta-analyses of binary outcomes from 1,991 Cochrane reviews, we calculated power per study within each meta-analysis. We defined adequate power as ≥50% power to detect a 30% relative risk reduction. In a subset of 1,107 meta-analyses including 5 or more studies with at least two adequately powered and at least one underpowered, results were compared with and without underpowered studies. In 10,492 (70%) of 14,886 meta-analyses, all included studies were underpowered; only 2,588 (17%) included at least two adequately powered studies. 34% of the meta-analyses themselves were adequately powered. The median of summary relative risks was 0.75 across all meta-analyses (inter-quartile range 0.55 to 0.89). In the subset examined, odds ratios in underpowered studies were 15% lower (95% CI 11% to 18%, $P<0.0001$) than in adequately powered studies, in meta-analyses of controlled pharmacological trials; and 12% lower (95% CI 7% to 17%, $P<0.0001$) in meta-analyses of controlled non-pharmacological trials. The standard error of the intervention effect increased by a median of 11% (inter-quartile range −1% to 35%) when underpowered studies were omitted; and between-study heterogeneity tended to decrease.

Conclusions: When at least two adequately powered studies are available in meta-analyses reported by Cochrane reviews, underpowered studies often contribute little information, and could be left out if a rapid review of the evidence is required. However, underpowered studies made up the entirety of the evidence in most Cochrane reviews.

BMJ

RESEARCH

Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study

Eveline Nüesch, research fellow,[1,2] Sven Trelle, associate director,[1,2] Stephan Reichenbach, senior research fellow,[1,3] Anne W S Rutjes, senior research fellow,[1,4] Beatrice Tschannen, research fellow,[1] Douglas G Altman, director and professor of statistics in medicine,[5] Matthias Egger, head of department and professor of epidemiology and public health,[1] Peter Jüni, head of division and professor of clinical epidemiology[1,2]

[1]Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland
[2]CTU Bern, Bern University Hospital, Switzerland
[3]Department of Rheumatology, Immunology and Allergology, Bern University Hospital, Switzerland
[4]Laboratory of Clinical Epidemiology of Cardiovascular Disease, Department of Clinical Pharmacology and Epidemiology, Consorzio Mario Negri Sud, Santa Maria Imbaro, Italy.
[5]Centre for Statistics in Medicine, University of Oxford, Oxford

Correspondence to: P Jüni
juni@ispm.unibe.ch

Cite this as: *BMJ* 2010;341:c3515
doi:10.1136/bmj.c3515

ABSTRACT

Objective To examine the presence and extent of small study effects in clinical osteoarthritis research.

Design Meta-epidemiological study.

Data sources 13 meta-analyses including 153 randomised trials (41 605 patients) that compared therapeutic interventions with placebo or non-intervention control in patients with osteoarthritis of the hip or knee and used patients' reported pain as an outcome.

Methods We compared estimated benefits of treatment between large trials (at least 100 patients per arm) and small trials, explored funnel plots supplemented with lines of predicted effects and contours of significance, and used three approaches to estimate treatment effects: meta-analyses including all trials irrespective of sample size, meta-analyses restricted to large trials, and treatment effects predicted for large trials.

Results On average, treatment effects were more beneficial in small than in large trials (difference in effect sizes −0.21, 95% confidence interval −0.34 to −0.08, P=0.001). Depending on criteria used, six to eight funnel plots indicated small study effects. In six of 13 meta-analyses, the overall pooled estimate suggested a clinically relevant, significant benefit of treatment, whereas analyses restricted to large trials and predicted effects in large trials yielded smaller non-significant estimates.

Conclusions Small study effects can often distort results of meta-analyses. The influence of small trials on estimated treatment effects should be routinely assessed.

selecting patients and implementing the experimental intervention.[9] The funnel plot is a scatter plot of treatment effects against standard error as a measure of statistical precision.[9 10] Imprecision of estimated treatment effects will increase as the sample size of component trials decreases. Thus, in the absence of small study effects, results from small trials with large standard errors will scatter widely at the bottom of a funnel plot while the spread narrows with increasing sample size and the plot will resemble a symmetrical inverted funnel. Conversely, if small study effects are present, funnel plots will be asymmetrical.[9] The plot can be enhanced by lines of the predicted treatment effect from meta-regression with the standard error as explanatory variable[11 12] and contours that divide the plot into areas of significance and non-significance.[13 14] A recent study of trials of anti-depressants[15] found that these approaches increased the understanding of the interplay of several biases associated with small sample size, including publication bias, selective reporting of outcomes, and inadequate methods and analysis.[14]

Small study effects are not uncommon in osteoarthritis research; several recent meta-analyses found pronounced asymmetry of funnel plots.[16-18] We previously studied the influence of methodological characteristics on estimated effects in a set of clinical osteoarthritis trials that used pain outcomes reported by patients and found that deficiencies in concealment of random allocation, blinding of patients, and analyses can distort the results in these trials.[19 20] Different components of inadequate trial methods often concur. A trial with adequate allocation concealment, for

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 67 (2014) 984–992

Journal of Clinical Epidemiology

Bias and small-study effects influence treatment effect estimates: a meta-epidemiological study in oral medicine

Spyridon N. Papageorgiou[a,b,c,*], Georgios N. Antonoglou[d], Elli Tsiranidou[e], Søren Jepsen[c,f], Andreas Jäger[a,c]

[a]*Department of Orthodontics, School of Dentistry, University of Bonn, Welschnonnenstr. 17, 53111, Bonn, Germany*
[b]*Department of Oral Technology, School of Dentistry, University of Bonn, Welschnonnenstr. 17, 53111, Bonn, Germany*
[c]*Clinical Research Unit 208, University of Bonn, 53111, Bonn, Germany*
[d]*Department of Periodontology and Geriatric Dentistry, Institute of Dentistry, University of Oulu, Medical Campus, Aapistie 3A, PO 5281, 90014, Oulu, Finland*
[e]*Medical Research Methodology, Medical School, Aristotle University, PO 17034, 54210, Thessaloniki, Greece*
[f]*Department of Periodontology, Operative and Preventive Dentistry, School of Dentistry, University of Bonn, Welschnonnenstr. 17, 53111, Bonn, Germany*

Accepted 4 April 2014; Published online 21 May 2014

Abstract

Objectives: To examine the influence of the following study characteristics on their study effect estimates: (1) indexing in MEDLINE, (2) language, and (3) design. For randomized trials, (4) trial size and (5) unequal randomization were also assessed.

Study Design and Setting: The CAtegorical Dental and Maxillofacial Outcome Syntheses meta-epidemiologic study was conducted. Eight databases/registers were searched up to September 2012 for meta-analyses of binary outcomes with at least five studies in the field of dental and maxillofacial medicine. The previously mentioned five study characteristics were investigated. The ratio of odds ratios (ROR) according to each characteristic was calculated with random-effects meta-regression and then pooled across meta-analyses.

Results: A total of 281 meta-analyses were identified and used to assess the influence of the following factors: non-MEDLINE indexing vs. MEDLINE indexing ($n = 78$; ROR, 1.12; 95% confidence interval [CI]: 1.05, 1.19; $P = 0.001$), language ($n = 61$; $P = 0.546$), design ($n = 24$; $P = 0.576$), small trials (<200 patients) vs. large trials (≥200 patients) ($n = 80$; ROR, 0.92; 95% CI: 0.87, 0.98; $P = 0.009$) and unequal randomization ($n = 36$; $P = 0.828$).

Conclusion: Studies indexed in MEDLINE might present greater effects than non-indexed ones. Small randomized trials might present greater effects than large ones.

Keywords: Dentistry; Meta-analysis; Systematic review; Effect size; Meta-epidemiologic study; systematic error

Zhang *et al. Critical Care* 2013, **17**:R2
http://ccforum.com/content/17/1/R2

CRITICAL CARE

RESEARCH Open Access

Small studies may overestimate the effect sizes in critical care meta-analyses: a meta-epidemiological study

Zhongheng Zhang*, Xiao Xu and Hongying Ni

Abstract

Introduction: Small-study effects refer to the fact that trials with limited sample sizes are more likely to report larger beneficial effects than large trials. However, this has never been investigated in critical care medicine. Thus, the present study aimed to examine the presence and extent of small-study effects in critical care medicine.

Methods: Critical care meta-analyses involving randomized controlled trials and reported mortality as an outcome measure were considered eligible for the study. Component trials were classified as large (≥100 patients per arm) and small (<100 patients per arm) according to their sample sizes. Ratio of odds ratio (ROR) was calculated for each meta-analysis and then RORs were combined using a meta-analytic approach. ROR<1 indicated larger beneficial effect in small trials. Small and large trials were compared in methodological qualities including sequence generating, blinding, allocation concealment, intention to treat and sample size calculation.

Results: A total of 27 critical care meta-analyses involving 317 trials were included. Of them, five meta-analyses showed statistically significant RORs <1, and other meta-analyses did not reach a statistical significance. Overall, the pooled ROR was 0.60 (95% CI: 0.53 to 0.68); the heterogeneity was moderate with an I^2 of 50.3% (chi-squared = 52.30; $P = 0.002$). Large trials showed significantly better reporting quality than small trials in terms of sequence generating, allocation concealment, blinding, intention to treat, sample size calculation and incomplete follow-up data.

Conclusions: Small trials are more likely to report larger beneficial effects than large trials in critical care medicine, which could be partly explained by the lower methodological quality in small trials. Caution should be practiced in the interpretation of meta-analyses involving small trials.

FEATURE ARTICLE

Publication Bias & Small-study Effects in Pediatric Dentistry Meta-analyses

Spyridon N. Papageorgiou[a,b,c], Dionysia Dimitraki[d], Trilby Coolidge[e], and Nikolaos Kotsanos[d]

[a]Department of Orthodontics, School of Dentistry, University of Bonn, Bonn, Germany
[b]Department of Oral Technology, School of Dentistry, University of Bonn, Bonn, Germany
[c]Clinical Research Unit 208, University of Bonn, Bonn, Germany
[d]Department of Paediatric Dentistry, School of Dentistry, Aristotle University of Thessaloniki, Thessaloniki, Greece
[e]Department of Oral Health Sciences, School of Dentistry, University of Washington, Seattle, WA, USA

Abstract

Objectives: The aim of this study was to examine the presence and extent of publication bias and small-study effects in meta-analyses (MAs) investigating pediatric dentistry-related subjects.

Methods: Following a literature search, 46 MAs including 882 studies were analyzed qualitatively. Of these, 39 provided enough data to be re-analyzed. Publication bias was assessed with the following methods: contour-enhanced funnel plots, Begg and Mazumdar's rank correlation and Egger's linear regression tests, Rosenthal's failsafe N, and Duval and Tweedie's "trim and fill" procedure.

Results: Only a few MAs adequately assessed the existence and effect of publication bias. Inspection of the funnel plots indicated asymmetry, which was confirmed by Begg–Mazumdar's test in 18% and by Egger's test in 33% of the MAs. According to Rosenthal's criterion, 80% of the MAs were robust, while adjusted effects with unpublished studies differed from little to great from the unadjusted ones. Pooling of the Egger's intercepts indicated that evidence of asymmetry was found in the pediatric dental literature, which was accentuated in dental journals and in diagnostic MAs. Since indications of small-study effects and publication bias in pediatric dentistry were found, the influence of small or missing trials on estimated treatment effects should be routinely assessed in future MAs.

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 68 (2015) 860–869

Journal of Clinical Epidemiology

Small studies are more heterogeneous than large ones: a meta-meta-analysis

Joanna IntHout[a,*], John P.A. Ioannidis[b,c,d,e], George F. Borm[a], Jelle J. Goeman[a]

[a]*Radboud Institute for Health Sciences (RIHS), Radboud University Medical Center, Mailbox 133, P.O. Box 9101, Nijmegen 6500 HB, The Netherlands*
[b]*Stanford Prevention Research Center, Department of Medicine, Stanford University School of Humanities and Sciences, 1265 Welch Road, Stanford, CA 94305, USA*
[c]*Department of Health Research and Policy, Stanford University School of Medicine, 150 Governor's Lane, Stanford, CA 94305, USA*
[d]*Department of Statistics, Stanford University School of Humanities and Sciences, 390 Serra Mall, Stanford, CA 94305, USA*
[e]*Meta-Research Innovation Center at Stanford (METRICS), Stanford University, 1265 Welch Road, Stanford, CA 94305, USA*

Accepted 27 March 2015; Published online 2 April 2015

Abstract

Objectives: Between-study heterogeneity plays an important role in random-effects models for meta-analysis. Most clinical trials are small, and small trials are often associated with larger effect sizes. We empirically evaluated whether there is also a relationship between trial size and heterogeneity (τ).

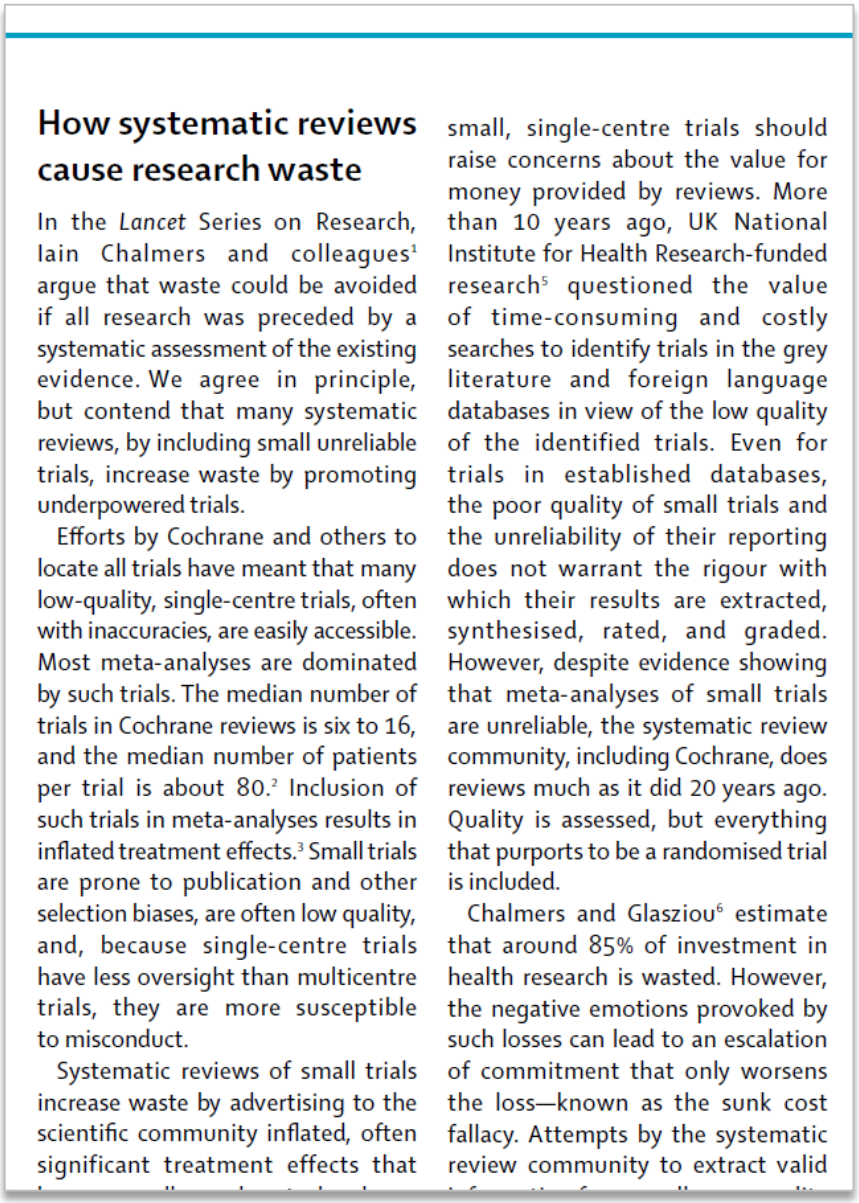
Study Design and Setting: We selected the first meta-analysis per intervention review of the Cochrane Database of Systematic Reviews Issues 2009–2013 with a dichotomous (n = 2,009) or continuous (n = 1,254) outcome. The association between estimated τ and trial size was evaluated across meta-analyses using regression and within meta-analyses using a Bayesian approach. Small trials were predefined as those having standard errors (SEs) over 0.2 standardized effects.

Results: Most meta-analyses were based on few (median 4) trials. Within the same meta-analysis, the small study τ_S^2 was larger than the large-study τ_L^2 [average ratio 2.11; 95% credible interval (1.05, 3.87) for dichotomous and 3.11 (2.00, 4.78) for continuous meta-analyses]. The imprecision of τ_S was larger than of τ_L: median SE 0.39 vs. 0.20 for dichotomous and 0.22 vs. 0.13 for continuous small-study and large-study meta-analyses.

Conclusion: Heterogeneity between small studies is larger than between larger studies. The large imprecision with which τ is estimated in a typical small-studies' meta-analysis is another reason for concern, and sensitivity analyses are recommended.

Keywords: Randomized controlled trial; Meta-analysis; Between-study heterogeneity; Random-effects model; Trial size; Cochrane Database of systematic reviews (CDSR)

University of BRISTOL



How systematic reviews cause research waste

In the *Lancet* Series on Research, Iain Chalmers and colleagues[1] argue that waste could be avoided if all research was preceded by a systematic assessment of the existing evidence. We agree in principle, but contend that many systematic reviews, by including small unreliable trials, increase waste by promoting underpowered trials.

Efforts by Cochrane and others to locate all trials have meant that many low-quality, single-centre trials, often with inaccuracies, are easily accessible. Most meta-analyses are dominated by such trials. The median number of trials in Cochrane reviews is six to 16, and the median number of patients per trial is about 80.[2] Inclusion of such trials in meta-analyses results in inflated treatment effects.[3] Small trials are prone to publication and other selection biases, are often low quality, and, because single-centre trials have less oversight than multicentre trials, they are more susceptible to misconduct.

Systematic reviews of small trials increase waste by advertising to the scientific community inflated, often significant treatment effects that small, single-centre trials should raise concerns about the value for money provided by reviews. More than 10 years ago, UK National Institute for Health Research-funded research[5] questioned the value of time-consuming and costly searches to identify trials in the grey literature and foreign language databases in view of the low quality of the identified trials. Even for trials in established databases, the poor quality of small trials and the unreliability of their reporting does not warrant the rigour with which their results are extracted, synthesised, rated, and graded. However, despite evidence showing that meta-analyses of small trials are unreliable, the systematic review community, including Cochrane, does reviews much as it did 20 years ago. Quality is assessed, but everything that purports to be a randomised trial is included.

Chalmers and Glasziou[6] estimate that around 85% of investment in health research is wasted. However, the negative emotions provoked by such losses can lead to an escalation of commitment that only worsens the loss—known as the sunk cost fallacy. Attempts by the systematic review community to extract valid

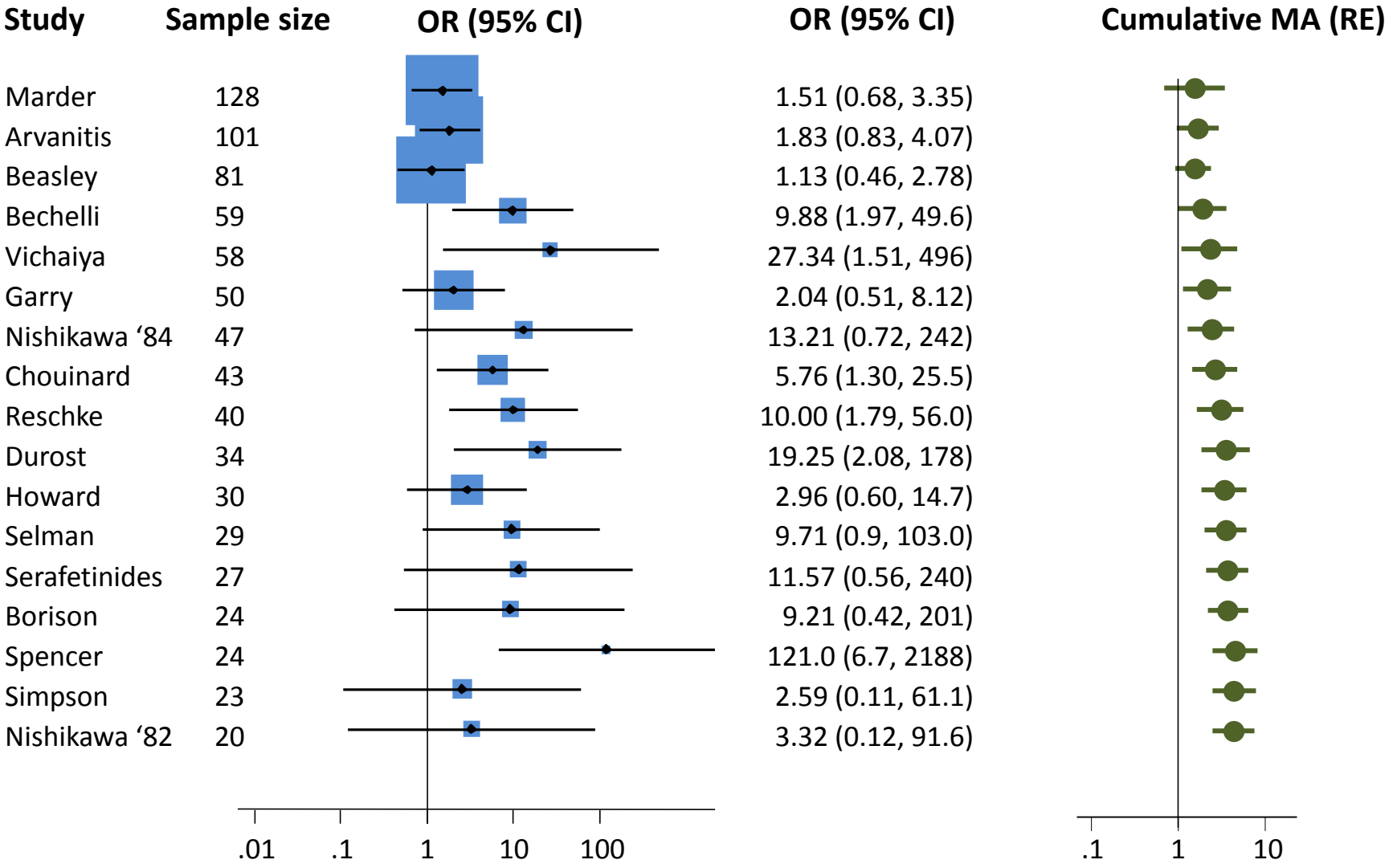
"Systematic reviews of small trials increase waste by advertising to the scientific community inflated, often significant treatment effects that become smaller or absent when large, high-quality trials are done"

"To ignore results from small trials and postulate plausible treatment effects that would be clinically worthwhile would be preferable."

Roberts and Kerr, Lancet 2015

University of BRISTOL

Cumulative forest plots



Clinical trials of haloperidol for schizophrenia

University of BRISTOL

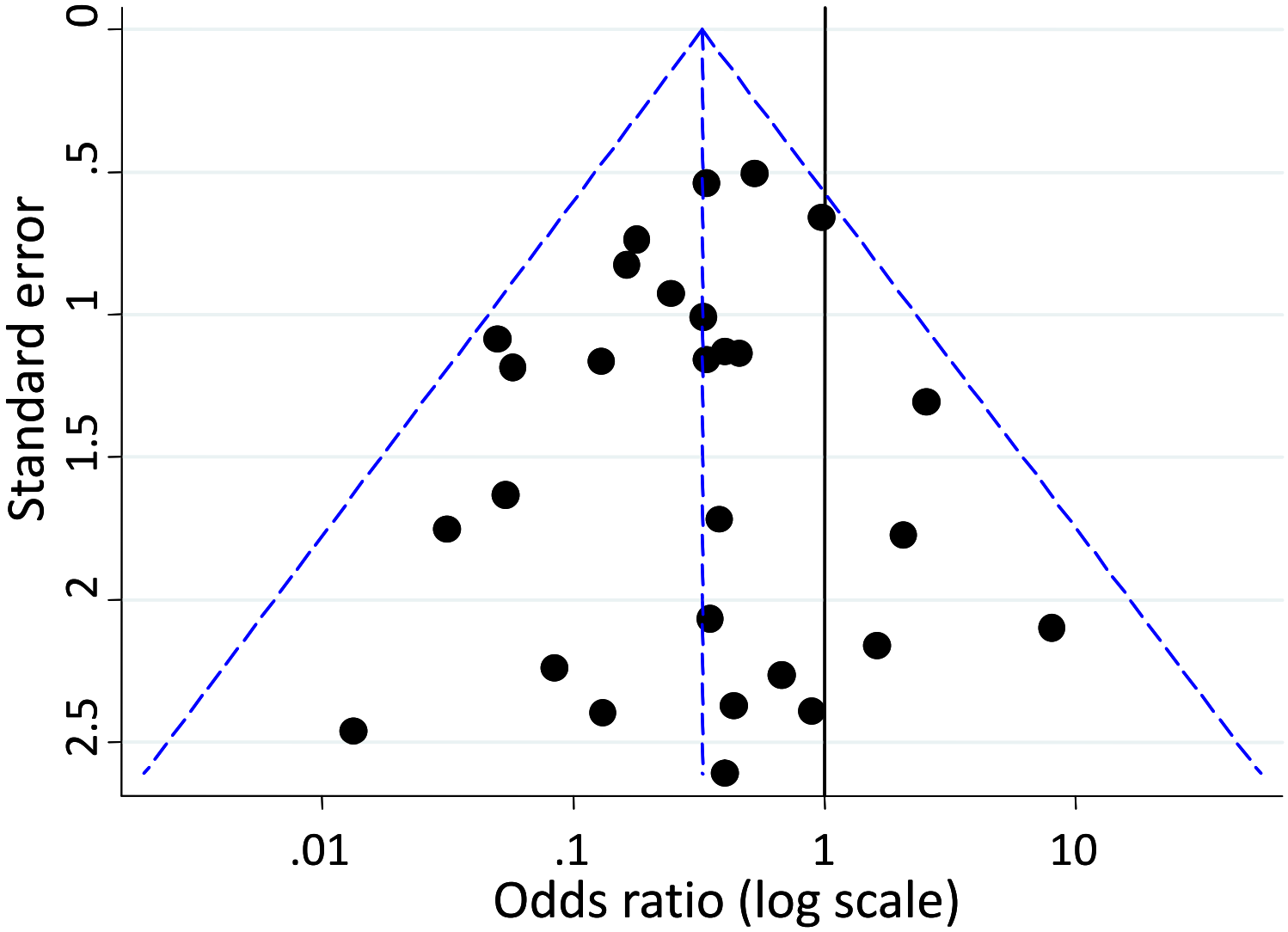
Sequential considerations

- Cumulative meta-analyses are not an inferential tool
- Repeated testing → False positive findings
- Sequential methods are available to control type I error

- Indeed, sequential methods offer one way to address small studies
  - Prevents early declarations of statistical significance based on the initial (potentially small) studies

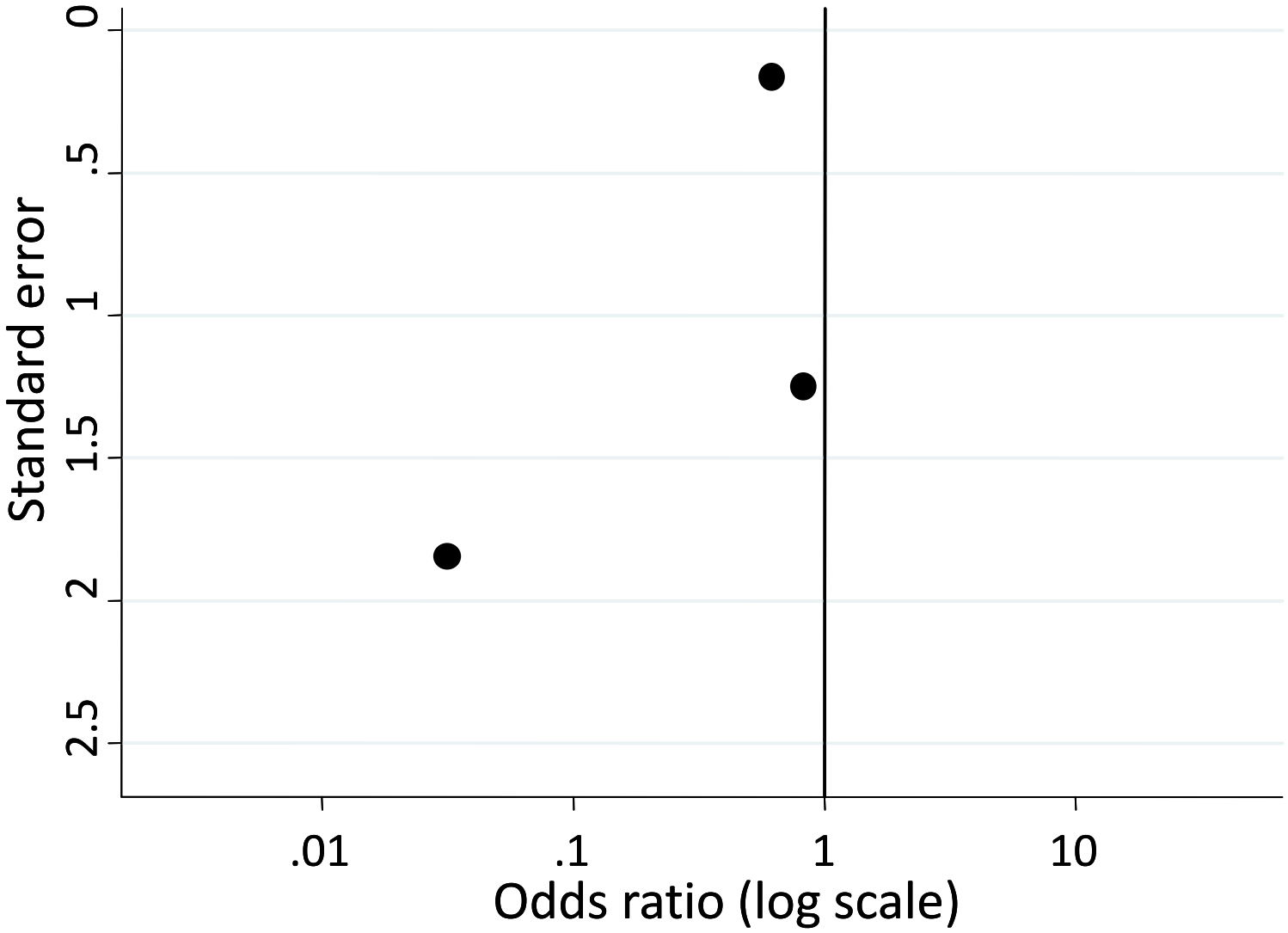
University of BRISTOL

Funnel plot

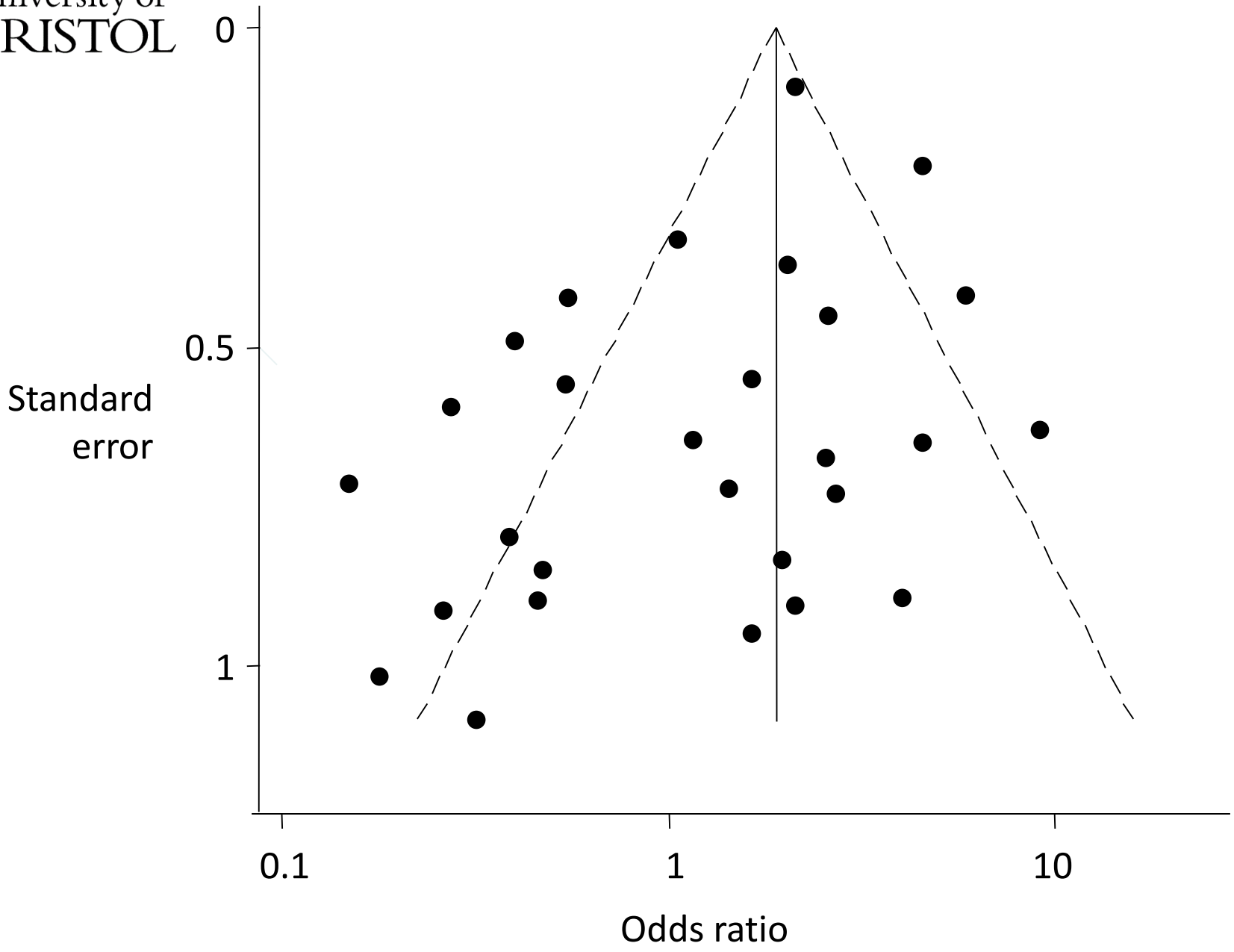


University of BRISTOL

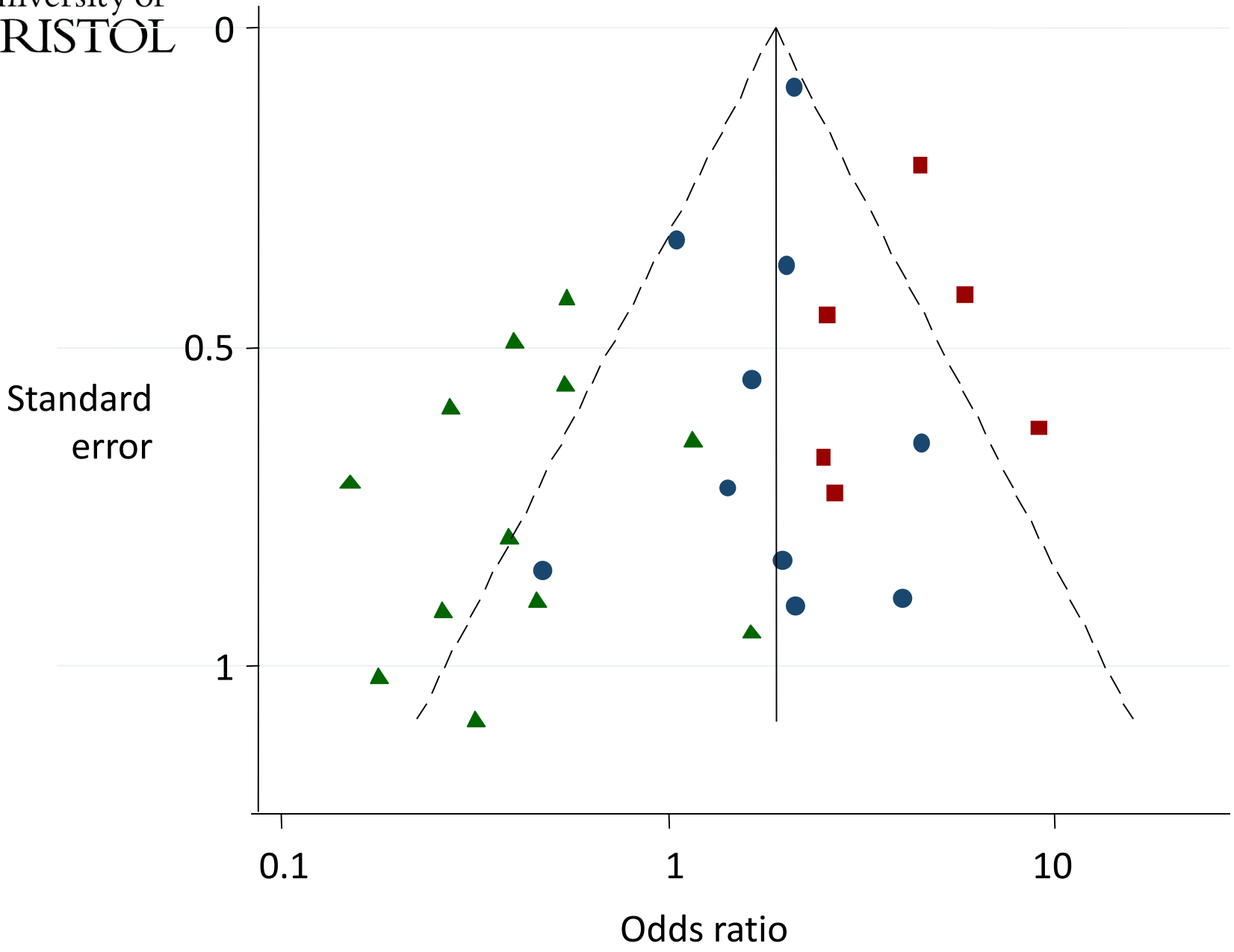
Funnel plot



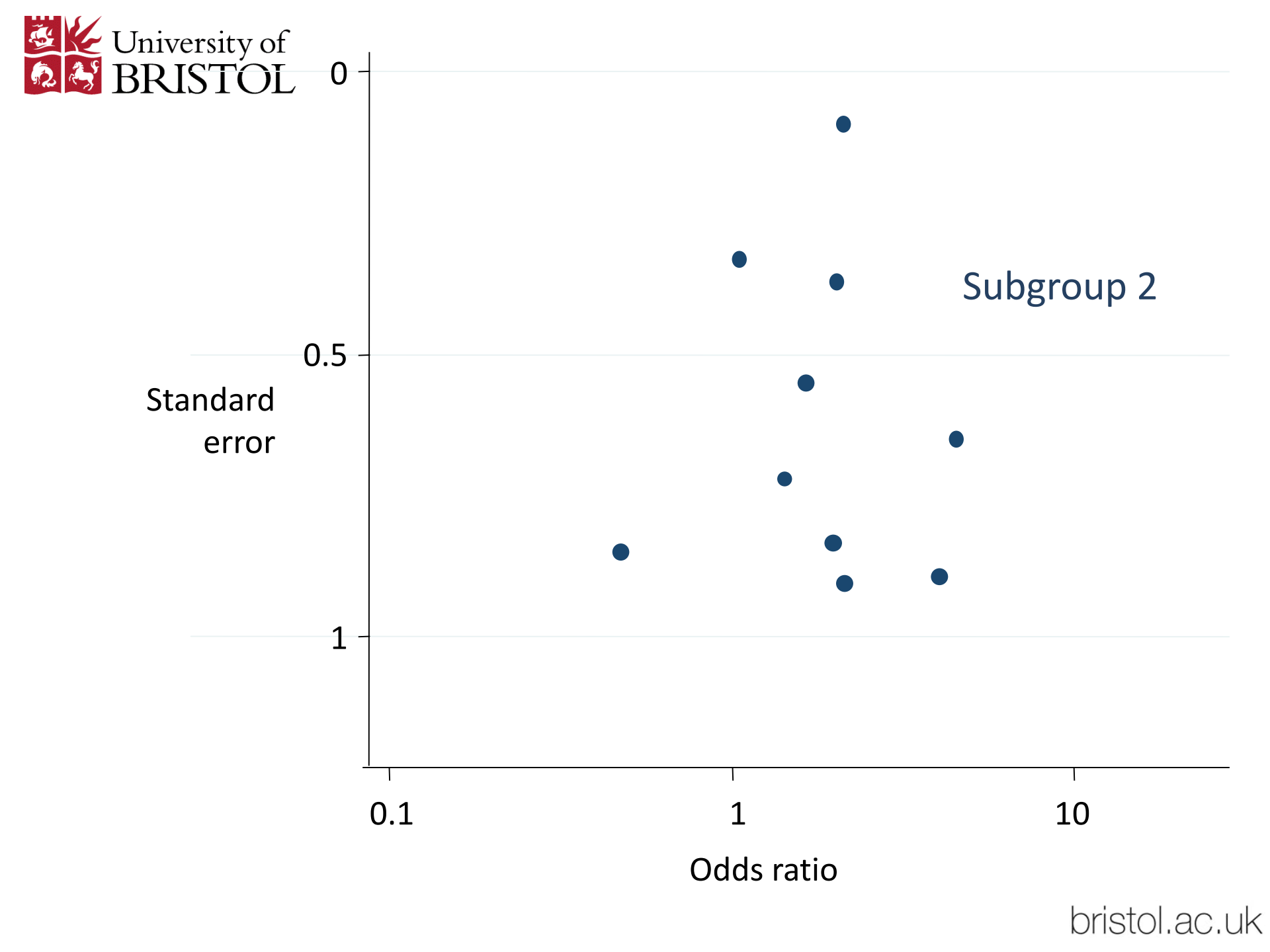
University of BRISTOL

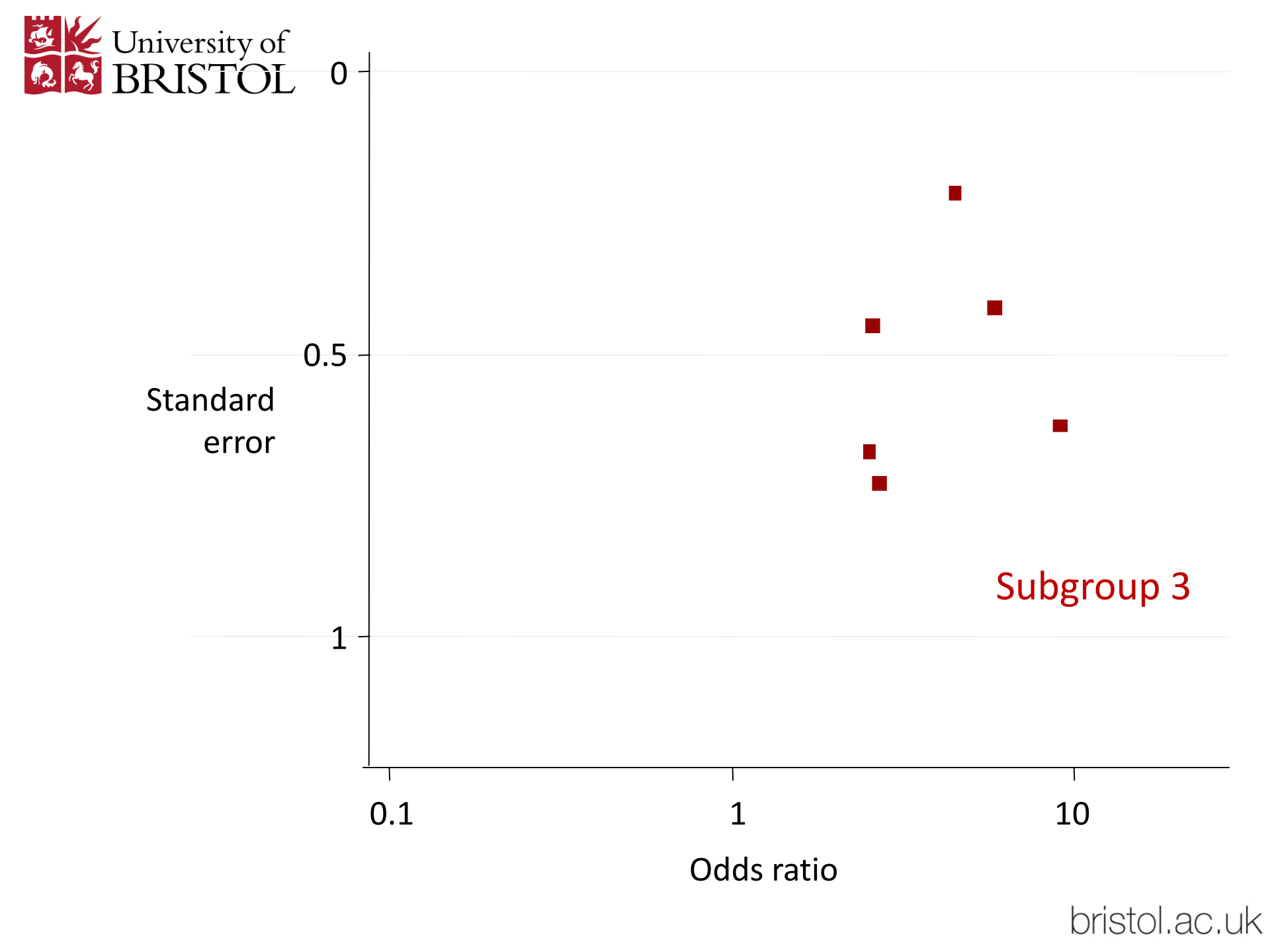


University of BRISTOL

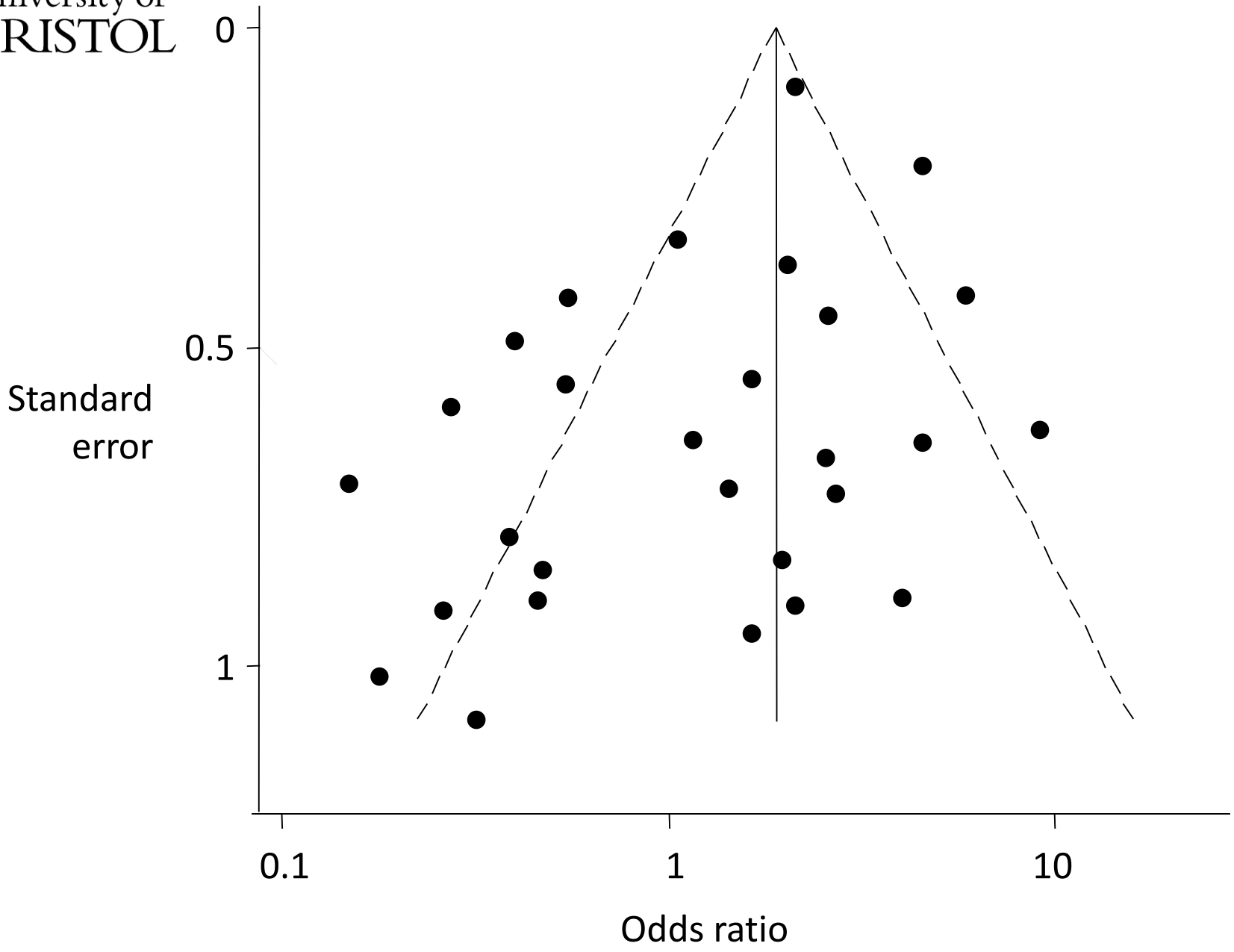



University of BRISTOL
0
0.5
1
Standard error
Subgroup 1
0.1
1
10
Odds ratio
bristol.ac.uk


University of BRISTOL
0
0.5
1
Standard error
Subgroup 2
0.1
1
10
Odds ratio
bristol.ac.uk


University of BRISTOL
0
0.5
1
Standard error
Subgroup 3
0.1
1
10
Odds ratio

University of BRISTOL



Box 1: Possible sources of asymmetry in funnel plots (adapted from Egger et al[1])

Reporting biases

- Publication bias:
  Delayed publication (also known as time lag or pipeline) bias

  Location biases (eg, language bias, citation bias, multiple publication bias)
- Selective outcome reporting
- Selective analysis reporting

Poor methodological quality leading to spuriously inflated effects in smaller studies

- Poor methodological design
- Inadequate analysis
- Fraud

True heterogeneity

- Size of effect differs according to study size (eg, because of differences in the intensity of interventions or in underlying risk between studies of different sizes)

Artefactual

- In some circumstances, sampling variation can lead to an association between the intervention effect and its standard error

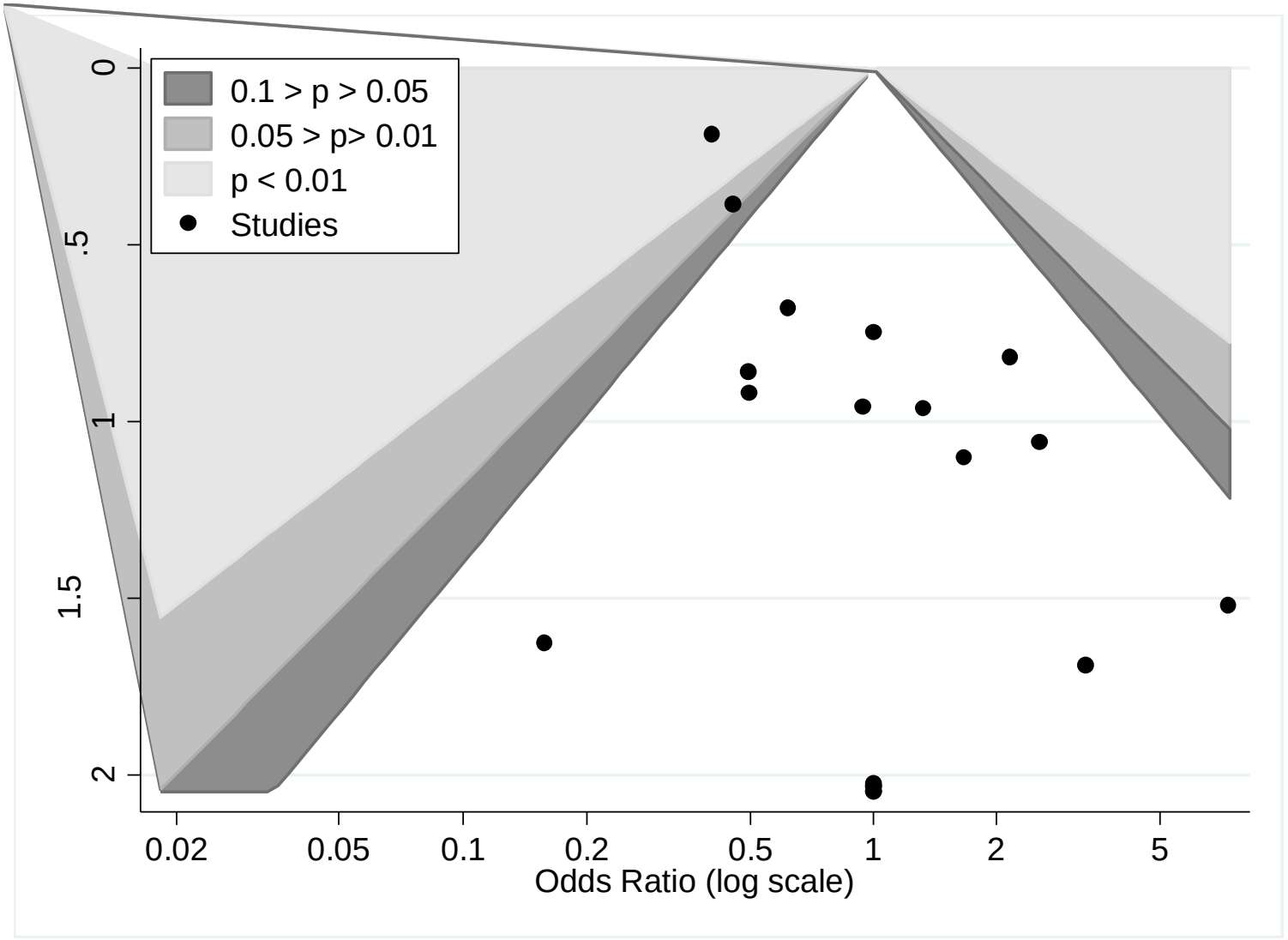
Chance

- Asymmetry may occur by chance, which motivates the use of asymmetry tests

Sterne et al (2011)

University of BRISTOL

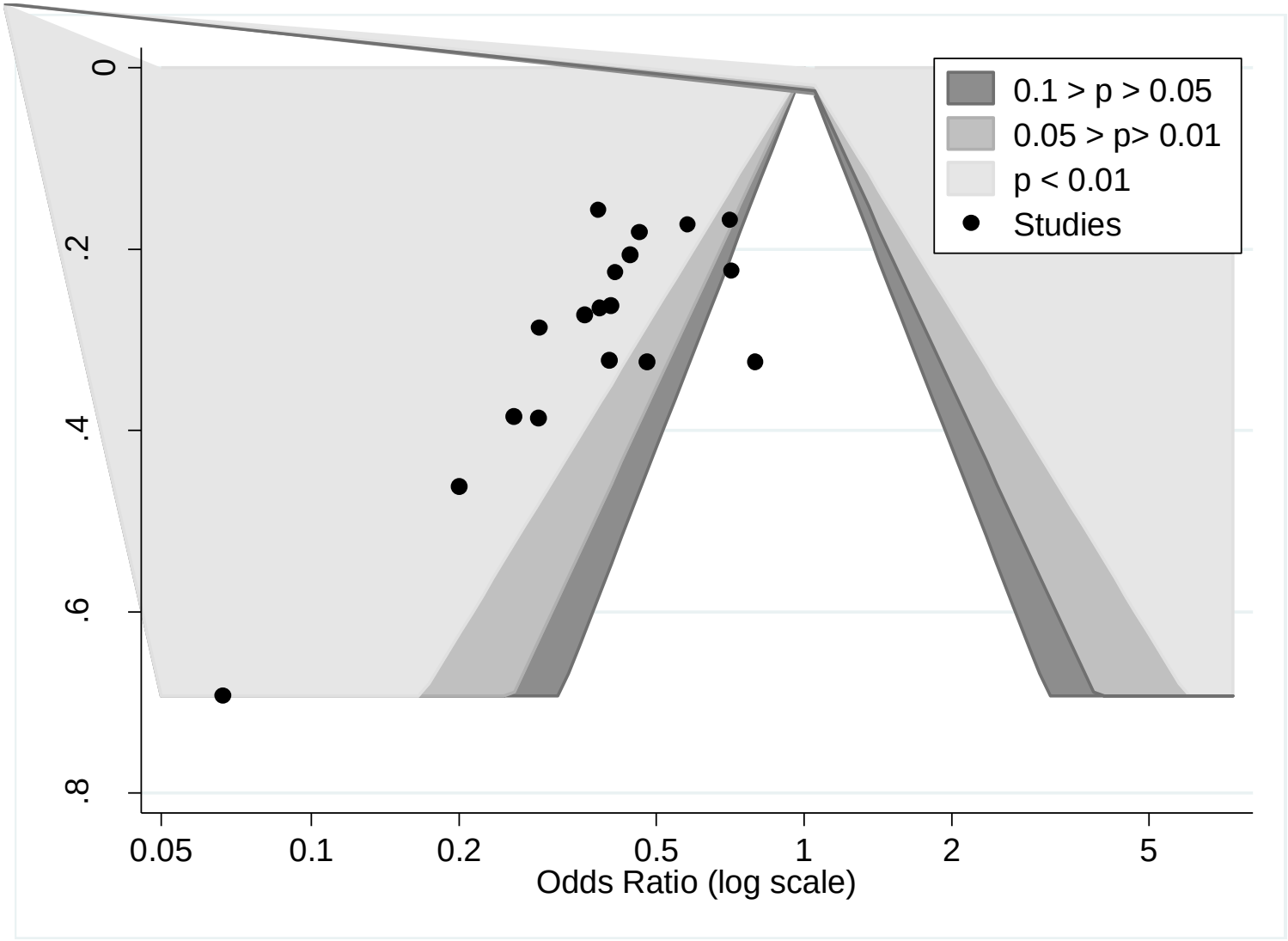
Contour-enhanced funnel plots



Peters et al (2008)

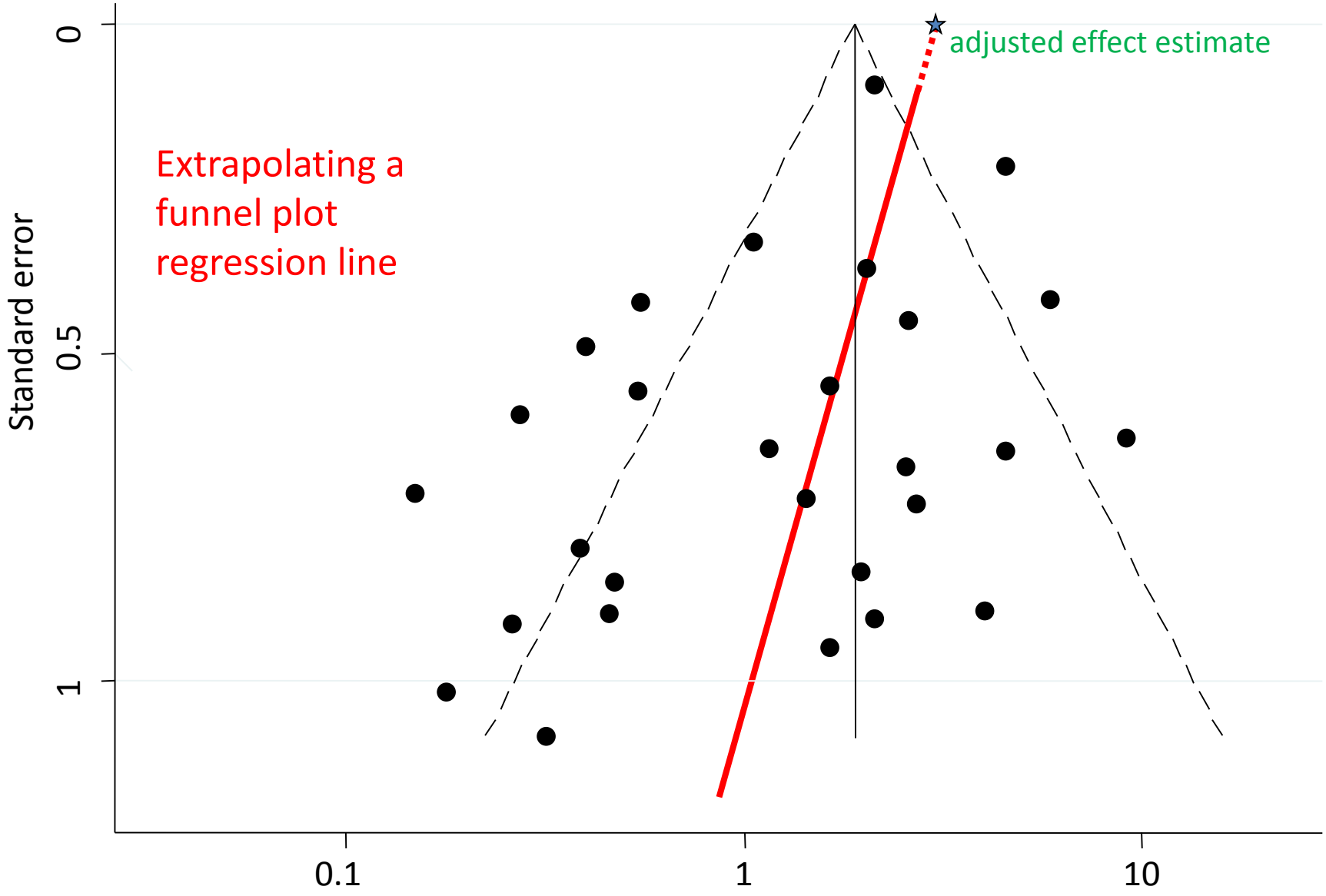
University of BRISTOL

Contour-enhanced funnel plots



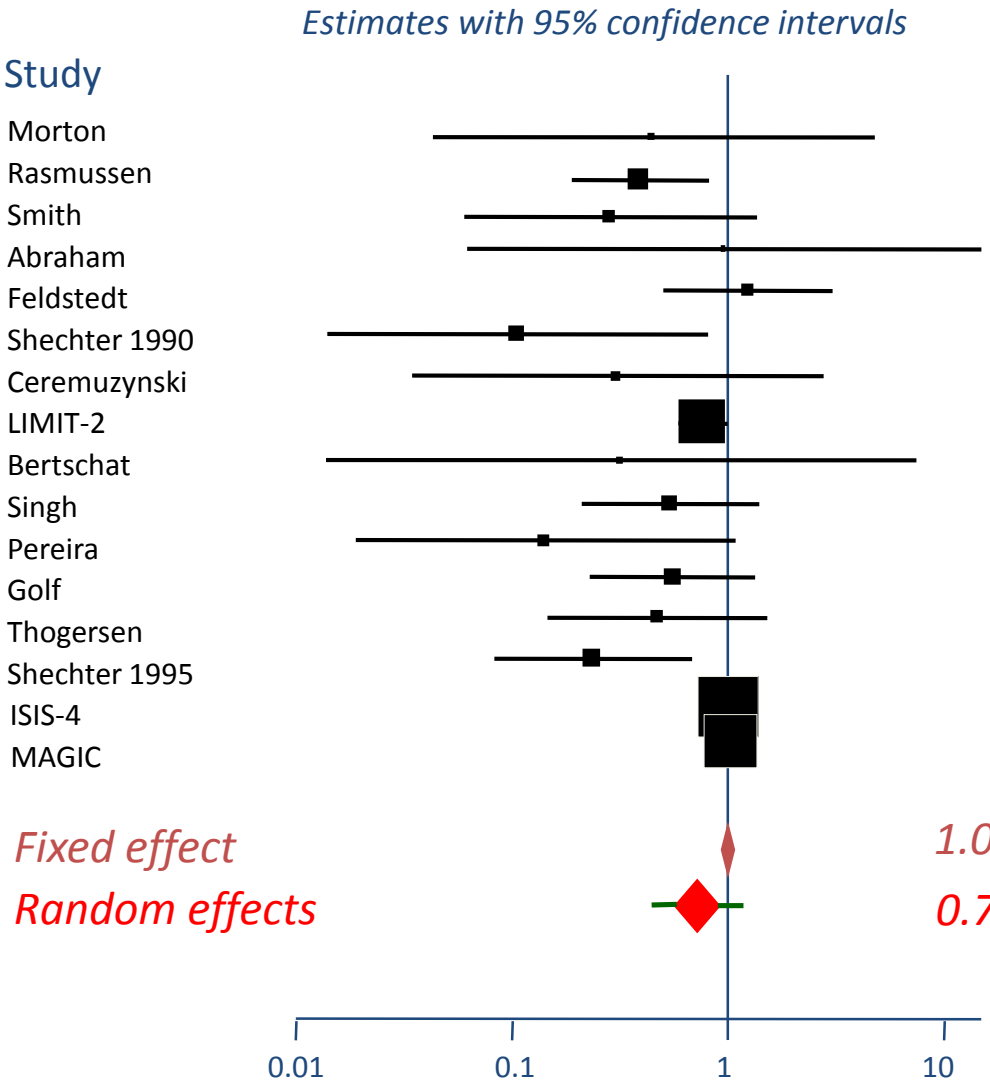
University of BRISTOL

Regression approach to small study effects



Moreno et al (2012)

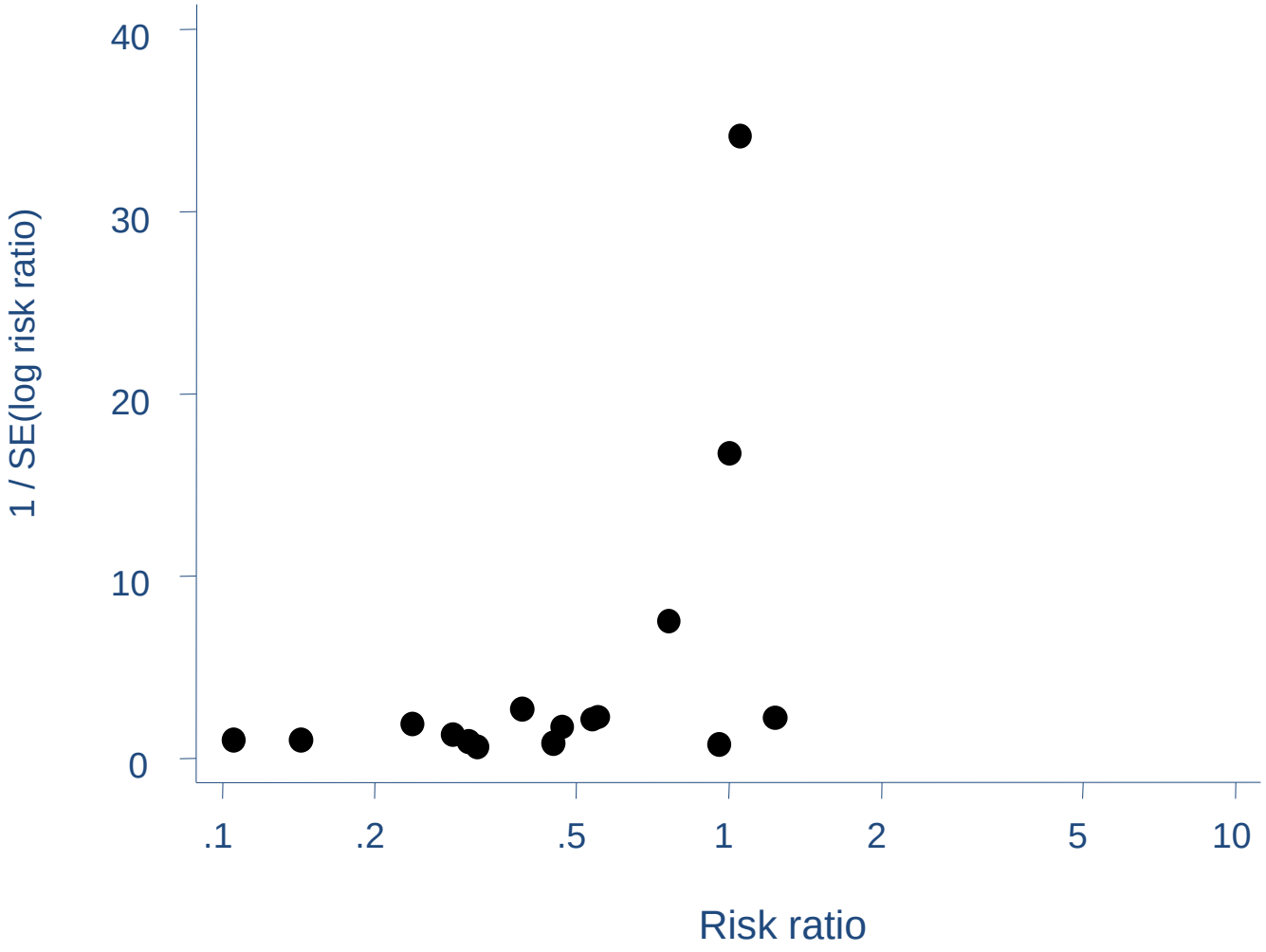
Magnesium trials



IV magnesium for acute MI (mortality)

University of BRISTOL

Funnel plot: asymmetrical



Meta-analyses that are more robust to small study effects (1)

- Henmi and Copas (2010) propose to use the **fixed-effect point estimate** but with variance that acknowledges heterogeneity.
- Given

$$y_i \sim N\left(\theta_i, \sigma_i^2\right)$$

$$\theta_i \sim N\left(\mu, \tau^2\right)$$

- then for any choice of ω_i (constants), μ is unbiasedly estimated by

$$\hat{\mu} = \frac{\sum \omega_i y_i}{\sum \omega_i}$$

- with standard error

$$\mathrm{Var}\left(\hat{\mu}\right) = \frac{\sum \omega_i^2 \left(\sigma_i^2 + \tau^2\right)}{\left(\sum \omega_i\right)^2}$$

Meta-analyses that are more robust to small study effects (2)

$$\mathrm{Var}(\hat{\mu}) = \frac{\sum \omega_i^2 (\sigma_i^2 + \tau^2)}{(\sum \omega_i)^2}$$

- Choosing weights $\omega_i = \frac{1}{\sigma_i^2}$

- gives usual fixed-effect estimate, with variance

$$\mathrm{Var}(\hat{\mu}) = \frac{\sum \left(\frac{\sigma_i^2 + \tau^2}{\sigma_i^4} \right)}{\left(\sum \frac{1}{\sigma_i^2} \right)^2}$$

- We could naively plug in estimates of τ and σ_i

- **Henmi and Copas** derive a confidence interval that accounts for uncertainty in τ (see their paper for R code)
- **Doi**'s *IVhet* meta-analysis is the same, but he uses the naive plug-in variance

University of BRISTOL

Some remarks on small *meta-analyses*

University of BRISTOL

Small meta-analyses

- When heterogeneity is present, random-effects meta-analysis may be appropriate

BUT

- Estimation of heterogeneity is difficult in small meta-analyses
- A descriptive analysis of Cochrane systematic reviews found that 75% of meta-analyses contained 5 or fewer studies (Davey et al., 2011)

- A **Bayesian approach** is very useful in small meta-analyses:
  - Allowance for all sources of uncertainty
  - Incorporation of external evidence

University of BRISTOL

Bayesian statistics

- We want to learn about some unknown quantities
  (e.g. odds ratio, mean difference, heterogeneity variance)

- A natural approach for accumulating data

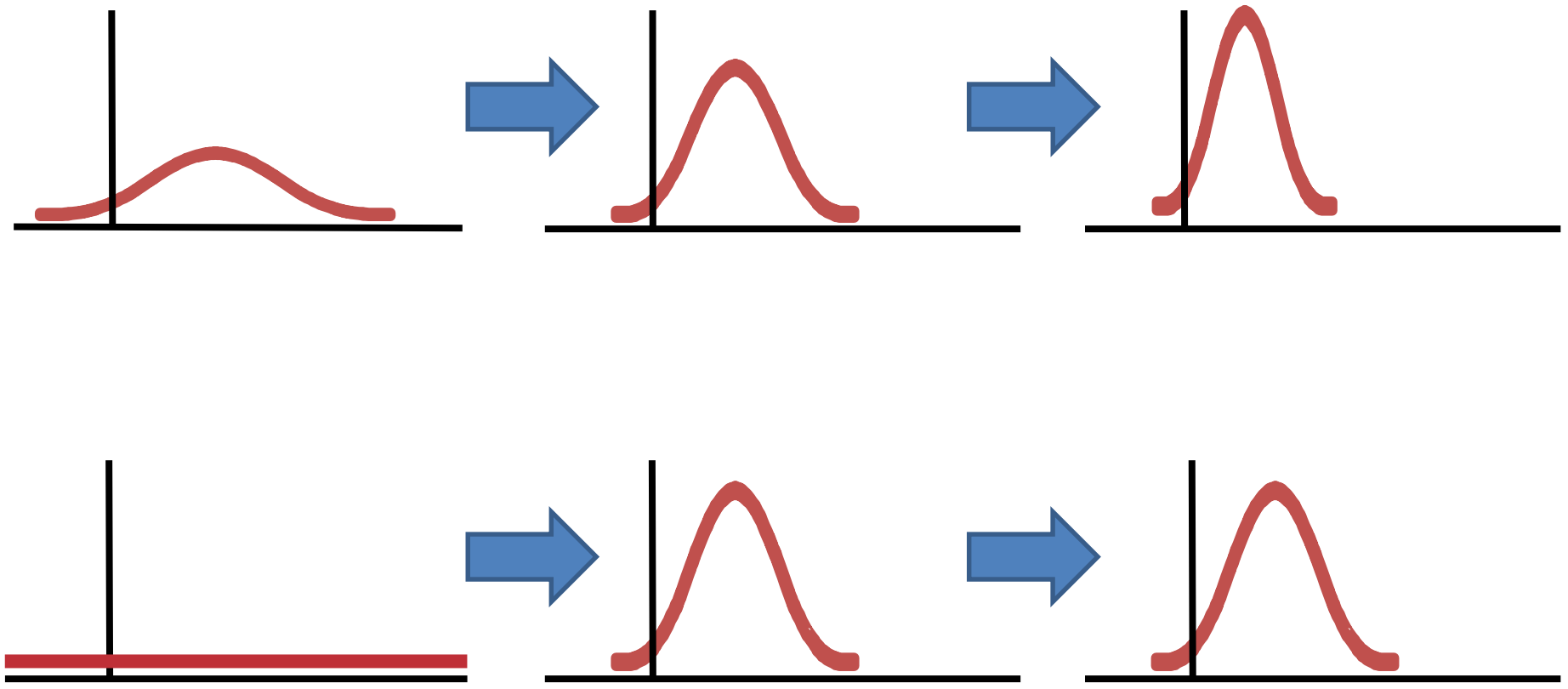
University of BRISTOL

Bayesian statistics

Prior

Likelihood

Posterior



University of BRISTOL

Davey et al. BMC Medical Research Methodology 2011, 11:160
http://www.biomedcentral.com/1471-2288/11/160

CORRESPONDENCE

Characteristics of

Predicting the extent meta-analysis, using *Cochrane Database o*

Rebecca M Turner,[1]* Jonathan Davey,[1] Mike

[1]MRC Biostatistics Unit, Institute of Public Health, Cambr
for Public Health, Queen's University Belfast, Northern Ire
University of Cambridge, Cambridge, UK

*Corresponding author. MRC Biostatistics Unit, Institute
E-mail: rebecca.turner@mrc-bsu.cam.ac.uk

Accepted 22 Februa

Background Many me
makes it
geneity.
evidence
random-e
ical evide
areas of

Methods Our anal
Database
according
and med
meta-ana
investigat
underlyin
butions
meta-ana

Results Between-
the outco
17% (95% CI 10–28) of variances for other
analyses comparing two active pharmacologica
erogeneity was on average 75% (95% CI 58–
non-pharmacological interventions. Meta-anal

ELSEVIER

CrossMark

Journal of Clinical Epidemiology 68 (2015) 52–60

Journal of Clinical Epidemiology

Predictive distributions were developed for the extent of heterogeneity in

[b,c]
s, Cambridge, CB2 0SR, UK
BS8 2PS, UK
DD, UK

meta-analysis is beneficial

Original Article

Research Synthesis Methods

Received 18 September 2014, Revised 5 November 2015, Accepted 6 November 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1193

Empirical evidence about inconsistency among st

Kirsty M. Rhodes,[a]*

**This paper investigates how inco
may differ, according to the type
data from 3873 binary, 5132
Database of Systematic Reviews.
were obtained, which can inform
on average for binary outcome
For a planned binary outcome
inconsistency among log odds ra
meta-analysis, the predictive d
median 40% and 95% CI: 15% to
odds ratios and log relative risk
meta-analyses using mean differ
The empirical evidence on incon
consistent in particular circumst
heterogeneity. © 2015 The Auth**

Statistics in Medicine

Research Article

Received 6 January 2014, Accepted 12 November 2014 Published online 5 December 2014 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6381

Predictive distributions for between-study heterogeneity and simple methods for their application in Bayesian meta-analysis

Rebecca M. Turner,[a*†] Dan Jackson,[a] Yinghui Wei,[b]
Simon G. Thompson[c] and Julian P. T. Higgins[d,e]

Numerous meta-analyses in healthcare research combine results from only a small number of studies, for which the variance representing between-study heterogeneity is estimated imprecisely. A Bayesian approach to estimation allows external evidence on the expected magnitude of heterogeneity to be incorporated.

The aim of this paper is to provide tools that improve the accessibility of Bayesian meta-analysis. We present two methods for implementing Bayesian meta-analysis, using numerical integration and importance sampling

University of BRISTOL

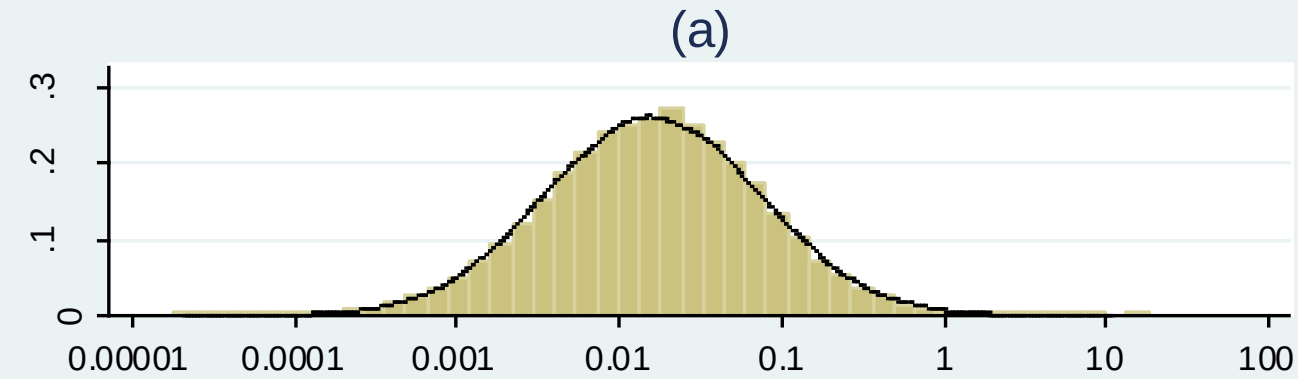
The idea

- Analyse lots of previous meta-analyses and look at how much heterogeneity there was
- Produce off-the-shelf predictive distributions for different types of meta-analyses
- These can be used as prior distributions for heterogeneity variance in new meta-analyses

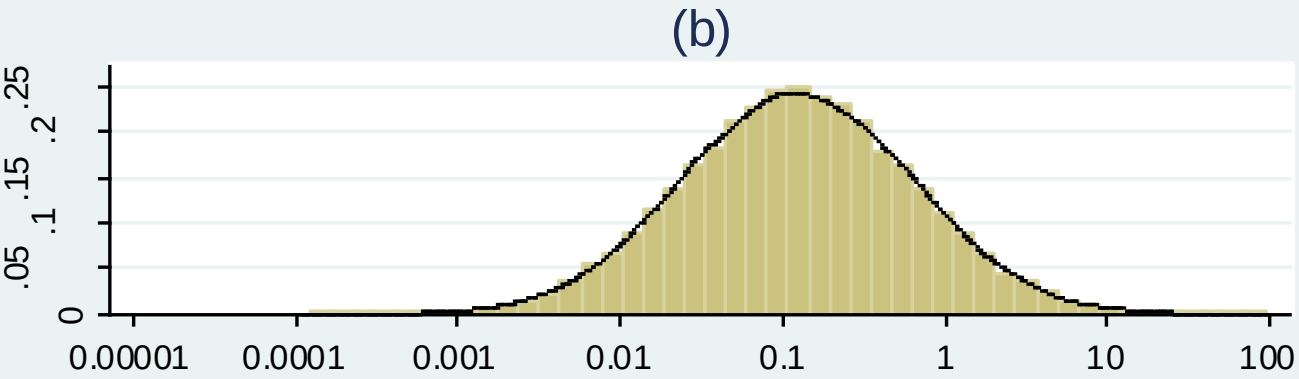
University of BRISTOL

Examples of predictive distributions for τ^2

Outcome: all-cause mortality; Comparison: Pharmacological vs. placebo/ctrl.



Outcome: subjective; Comparison: Non-pharma. (any comparator)

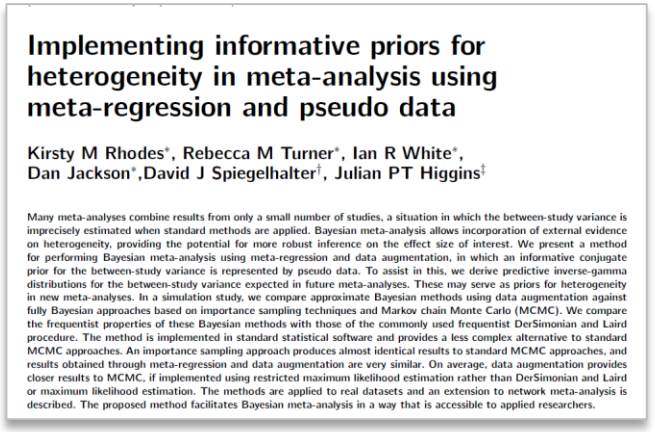


University of BRISTOL

Problem and solution

- Bayesian meta-analyses are computationally complex
- Usually done with simulation methods (Markov chain Monte Carlo) using WinBUGS or OpenBUGS

- An exciting recent development (Kirsty Rhodes et al, submitted) allows us to use informative prior distributions in Stata



Implementing informative priors for heterogeneity in meta-analysis using meta-regression and pseudo data

Kirsty M Rhodes*, Rebecca M Turner*, Ian R White*, Dan Jackson*, David J Spiegelhalter†, Julian PT Higgins‡

Many meta-analyses combine results from only a small number of studies, a situation in which the between-study variance is imprecisely estimated when standard methods are applied. Bayesian meta-analysis allows incorporation of external evidence on heterogeneity, providing the potential for more robust inference on the effect size of interest. We present a method for performing Bayesian meta-analysis using meta-regression and data augmentation, in which an informative conjugate prior for the between-study variance is represented by pseudo data. To assist in this, we derive predictive inverse-gamma distributions for the between-study variance expected in future meta-analyses. These may serve as priors for heterogeneity in new meta-analyses. In a simulation study, we compare approximate Bayesian methods using data augmentation against fully Bayesian approaches based on importance sampling techniques and Markov chain Monte Carlo (MCMC). We compare the frequentist properties of these Bayesian methods with those of the commonly used frequentist DerSimonian and Laird procedure. The method is implemented in standard statistical software and provides a less complex alternative to standard MCMC approaches. An importance sampling approach produces almost identical results to standard MCMC approaches, and results obtained through meta-regression and data augmentation are very similar. On average, data augmentation provides closer results to MCMC, if implemented using restricted maximum likelihood estimation rather than DerSimonian and Laird or maximum likelihood estimation. The methods are applied to real datasets and an extension to network meta-analysis is described. The proposed method facilitates Bayesian meta-analysis in a way that is accessible to applied researchers.

- We make up some fake studies and analyse them alongside the real data
- We use the fake studies to learn about the heterogeneity variance (but they don't contribute to treatment effect)

University of BRISTOL

Fake studies to reflect some prior distributions for heterogeneity variance

	Pharmacological vs Placebo/ Control	Pharmacological vs Pharmacological	Non-Pharmacological (Any)
All-cause mortality	$IG(1.06, 0.01)$ **Unobserved data:** 2 studies with effects $y_0 = 0.100$	$IG(2.93, 0.00003)$ **Unobserved data:** 6 studies with effects $y_0 = 0.003$	$IG(0.80, 0.007)$ **Unobserved data:** 2 studies with effects $y_0 = 0.084$
Semi-objective	$IG(1.32, 0.08)$ **Unobserved data:** 3 studies with effects $y_0 = 0.231$	$IG(1.04, 0.04)$ **Unobserved data:** 2 studies with effects $y_0 = 0.200$	$IG(0.88, 0.05)$ **Unobserved data:** 2 studies with effects $y_0 = 0.224$
Subjective	$IG(1.45, 0.18)$ **Unobserved data:** 3 studies with effects $y_0 = 0.346$	$IG(1.13, 0.09)$ **Unobserved data:** 2 studies with effects $y_0 = 0.300$	$IG(1.39, 0.13)$ **Unobserved data:** 3 studies with effects $y_0 = 0.294$

Stata

- We'll use the prior for subjective outcomes, pharmacol vs control
  - 3 studies with lnOR = 0.346 and very small standard error

14	Serafetinides	1972	4	10	0	0	13	1	13
15	Simpson	1967	2	14	0	0	7	1	14
16	Spencer	1992	11	1	0	1	11	0	10
17	Vichaiya	1971	9	20	1	0	29	1	12
18	fake	.	.	.	.	.	.	.	.

```
local new = _N+1
set obs `new'
replace author = "fake" in `new'
gen real=1
replace real=0 if author == "fake"
replace lnOR=0.346 if real == 0
replace se_lnOR=1E-10 if real == 0
expand 3 if real==0
metareg lnOR real, wsse(se_lnOR) reml z noconst eform
```

How many studies?
Create new study
Label as 'fake'
Dummy: real (1) vs fake (0)

Compute lnOR for fake
... and very small SE
Make it 3 fake studies

Regress on fake; no intercept

University of BRISTOL

Results for haloperidol

Method for tau	Method for confidence interval	Meta-analysis result (95% CI)	Estimate of τ^2
Fixed-effect analysis		OR = 2.85 (1.99, 4.10)	n/a
DerSimonian-Laird	Z (normal)	OR = 4.20 (2.42, 7.30)	0.48
DerSimonian-Laird	Hartung-Knapp (t)	OR = 4.20 (2.31, 7.64)	"
Paule-Mandel (empirical Bayes)	Z (normal)	OR = 4.14 (2.40, 7.13)	0.45
Paule-Mandel (empirical Bayes)	Hartung-Knapp (t)	OR = 4.14 (2.30, 7.45)	"
REML	Z (normal)	OR = 4.38 (2.44, 7.86)	0.60
REML	Hartung-Knapp (t)	OR = 4.38 (2.33, 8.24)	"
Profile likelihood	Z (normal)	OR = 4.25 (2.39, 8.82)	0.51
Random-effects logistic regression		OR = 4.72 (2.61, 8.53)	0.25
Bayesian analysis with prior		OR = 3.79 (2.32, 6.18)	0.27

University of BRISTOL

Concluding remarks

- Small studies pose problems
- They may have larger effect sizes (on average), which may be due to
  - within-study bias
  - reporting bias
  - heterogeneity
  - chance
- In principle it's more important to focus on bias than to implement differential policies for smaller and larger studies
  - although I recognize this is difficult in practice
- Random-effects logistic regression is available, and is probably the method we should always be using for binary data

University of BRISTOL

References

- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629-34.
- Henmi M, Copas JB. Confidence intervals for random effects meta-analysis and robustness to publication bias. *Statistics in Medicine* 2010; **29**: 2969-83.
- Moreno SG, Sutton AJ, Thompson JR, Ades AE, Abrams KR, Cooper NJ. A generalized weighting regression-derived meta-analysis estimator robust to small-study effects and heterogeneity. *Statistics in Medicine* 2012; **31**: 1407-17.
- Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. *Journal of Clinical Epidemiology* 2008; 61: 991-6. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry.
- Simmonds MC, Higgins JPT. A general framework for the use of logistic regression models in meta-analysis. *Statistical Methods in Medical Research*. Published online 12 May 2014.
- Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, Carpenter J, Rücker G, Harbord RM, Schmid CH, Tetzlaff J, Deeks JJ, Peters J, Macaskill P, Schwarzer G, Duval S, Altman DG, Moher D, Higgins JPT. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011; **343**: d4002
- Doi SA, Barendregt JJ, Khan S, Thalib L, Williams GM. Advances in the meta-analysis of heterogeneous clinical trials I: The inverse variance heterogeneity model. *Contemporary Clinical Trials* 2015; 45: 130-8.