

Introduction to Systematic Reviews & Meta-analyses

Madhukar Pai, MD, PhD

Assistant Professor

Department of Epidemiology & Biostatistics

McGill University, Montreal, Canada

Email: madhukar.pai@mcgill.ca



McGill

Case Study 1: “Egg on their faces: the story of human albumin solution”*

- Human albumin solution, a blood product, has been used in the treatment of blood loss and burns since the attack on Pearl Harbour over half a century ago.
- In the UK alone, an estimated 100,000 patients are treated with human albumin solution each year, at a cost to the NHS of close to 12 million.
- In 1996, the global albumin market was worth £900,000.
- But is human albumin administration beneficial?

*1. Roberts I, et al. Egg on their faces. The story of human albumin solution. Eval Health Prof. 2002;25(1):130-8.

2. Cochrane Injuries Group Albumin Reviewers. Human albumin administration in critically ill patients: systematic review of randomised controlled trials. BMJ 1998;317:235-40.

“Egg on their faces: the story of human albumin solution”

- To answer this question a systematic review of controlled trials comparing albumin with crystalloid was conducted by the Cochrane Injuries Group.
- 30 RCTs including 1419 randomised patients identified.
- A meta-analysis showed that the risk of death among those treated with albumin was higher than in the comparison groups.
- The pooled risk ratio was 1.68 (95% CI 1.26, 2.23)
- The data suggested that for every seventeen critically ill patients treated with albumin there is one extra death.

"Egg on their faces: the story of human albumin solution"

- "Despite vigorous attempts by the plasma products industry to limit the impact of the systematic review on albumin sales, the use of albumin declined steeply.
- Throughout the UK albumin sales fell by 40%.
- The decline in albumin use occurred despite vigorous criticism of the review in the letters pages of the BMJ.
- The decline in albumin sales is a clear indication that doctors took into account the evidence presented in the systematic review and that many doctors changed their practice in response."

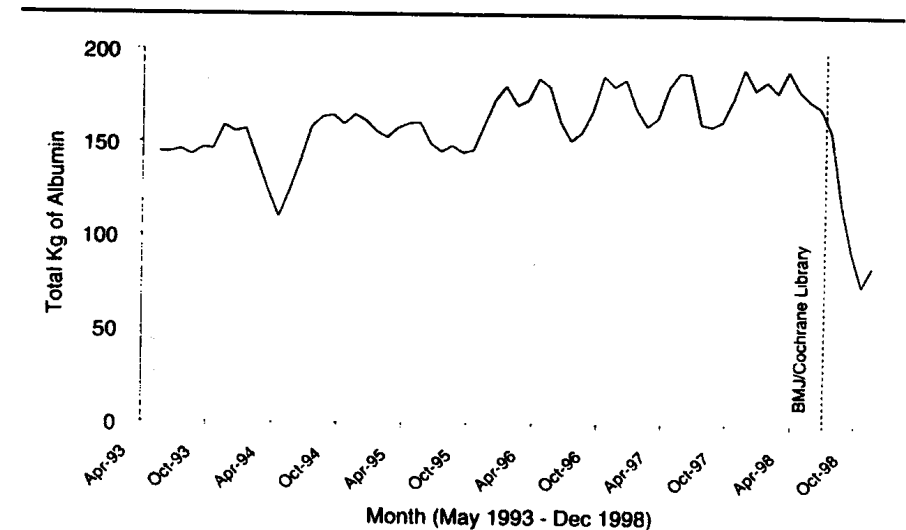


Figure 1: Albumin Sales in Scotland and Northern Ireland Before and After Publication of Systematic Review on Human Albumin Administration in Critically Ill Patients

Patient Survival after Human Albumin Administration

A Meta-Analysis of Randomized, Controlled Trials

Mahlon M. Wilkes, PhD, and Roberta J. Navickis, PhD

Purpose: To test the hypothesis that albumin administration is not associated with excess mortality.

Data Sources: Computer searches of the MEDLINE and EMBASE databases, the Cochrane Library, and Internet documents; hand searching of medical journals; inquiries to investigators and medical directors; and review of reference lists.

Study Selection: Randomized, controlled trials comparing albumin therapy with crystalloid therapy, no albumin, or lower doses of albumin.

Data Extraction: Two investigators independently extracted data. The primary end point was relative risk for death. Criteria used to assess methodologic quality were blinding, method of allocation concealment, presence of mortality as a study end point, and crossover. Small-trial bias was also investigated.

Data Synthesis: Fifty-five trials involving surgery or trauma,

burns, hypoalbuminemia, high-risk neonates, ascites, and other indications were included. Albumin administration did not significantly affect mortality in any category of indications. For all trials, the relative risk for death was 1.11 (95% CI, 0.95 to 1.28). Relative risk was lower among trials with blinding (0.73 [CI, 0.48 to 1.12]; $n = 7$), mortality as an end point (1.00 [CI, 0.84 to 1.18]; $n = 17$), no crossover (1.04 [CI, 0.89 to 1.22]; $n = 35$), and 100 or more patients (0.94 [CI, 0.77 to 1.14]; $n = 10$). In trials with two or more such attributes, relative risk was further reduced.

Conclusions: Overall, no effect of albumin on mortality was detected; any such effect may therefore be small. This finding supports the safety of albumin. The influence of methodologic quality on relative risk for death suggests the need for further well-designed clinical trials.

Ann Intern Med. 2001;135:149-164.

www.annals.org

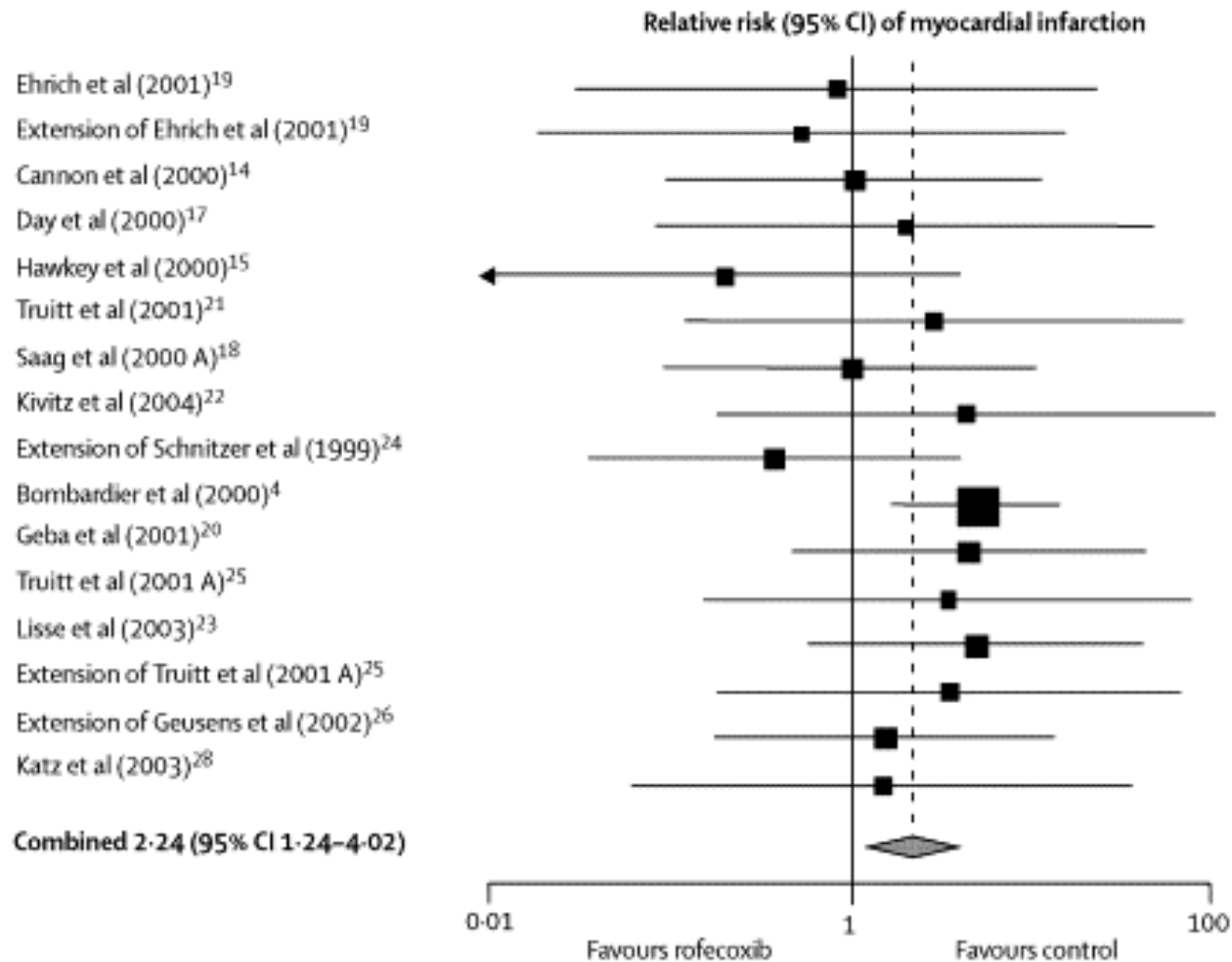
For author affiliations, current addresses, and contributions, see end of text.

See editorial comment on pp 205-208.

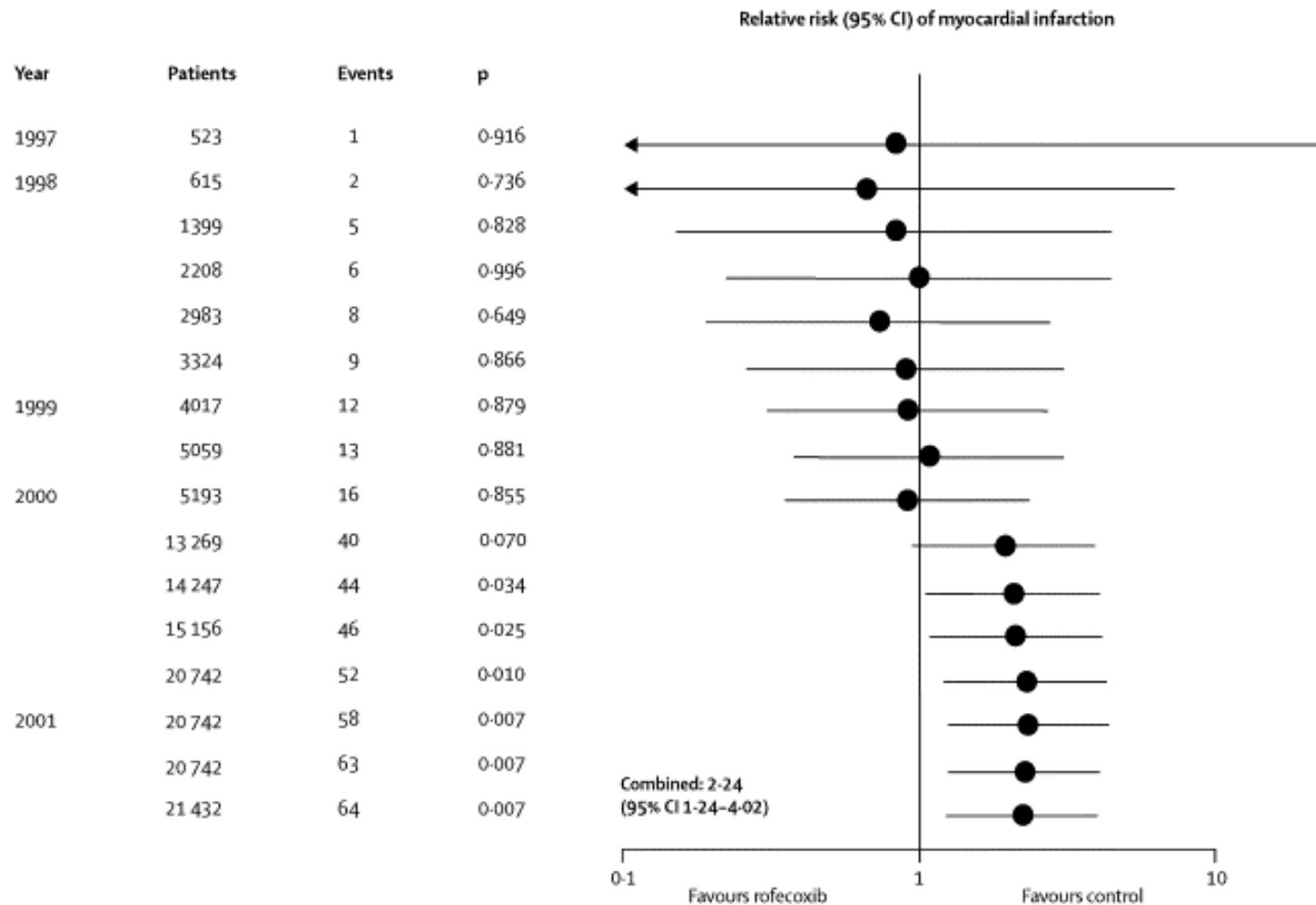
Case study 2: The Vioxx story

- On Sept 30, 2004, Merck announced the withdrawal of rofecoxib (Vioxx) because of an increased cardiovascular risk in patients taking the drug for >18 months
- Decision was based on the 3-year results of the unpublished APPROVe study, a RTC of rofecoxib for the prevention of colorectal polyps in patients with a history of colorectal adenomas
- By 2004, rofecoxib had been taken by ~ 80 million people (sales US\$2.5 billion)
- Juni et al. did a meta-analysis of 18 RCTs and 11 observational studies
- By the end of 2000 (52 events, 20742 patients) the relative risk from RTCs was 2.30 (95% CI 1.22–4.33, $p=0.010$), and 1 year later (64 events, 21432 patients) it was 2.24 (1.24–4.02, $p=0.007$).
- Juni et al. concluded that “rofecoxib should have been withdrawn several years earlier; the reasons why manufacturer and drug licensing authorities did not continuously monitor and summarise the accumulating evidence need to be clarified”

Case study 2: The Vioxx story



Case study 2: The Vioxx story



Case study 3. Does male circumcision reduce risk of HIV?

International Journal of STD & AIDS 1999; 10: 8–16

REVIEW ARTICLE

Circumcision and HIV infection: review of the literature and meta-analysis

R S Van Howe MD FAAP

Department of Pediatrics, Marshfield Clinic, Lakeland Center, USA

Summary: Thirty-five articles and a number of abstracts have been published in the medical literature looking at the relationship between male circumcision and HIV infection. Study designs have included geographical analysis, studies of high-risk patients, partner studies and random population surveys. Most of the studies have been conducted in Africa. A meta-analysis was performed on the 29 published articles where data were available. When the raw data are combined, a man with a circumcised penis is at greater risk of acquiring and transmitting HIV than a man with a non-circumcised penis (odds ratio (OR)=1.06, 95% confidence interval (CI)=1.01–1.12). Based on the studies published to date, recommending routine circumcision as a prophylactic measure to prevent HIV infection in Africa, or elsewhere, is scientifically unfounded.

Author	Intact+	Intact—	Circ+	Circ—	Odds ratio	95% CI	Relative risk	Attributable risk (%)
Truck drivers								
Bwayo ⁷	92	86	160	612	4.09	2.91–5.76	2.49	59.9
Bwayo ⁸	22	46	37	200	2.59	1.39–4.79	2.07	51.7
STD clinics								
Kreiss ²³	59	18	254	168	2.17	1.24–3.81	1.27	21.4
Hira ¹⁵	418	172	10	10	2.43	0.99–5.94	1.42	29.4
Cameron ¹¹	18	61	6	208	10.23	3.89–26.90	8.13	87.7
Pépin ^{44,14}	5	42	13	243	2.22	0.75–6.57	2.09	52.3
Greenblatt ¹²	11	28	8	68	3.34	1.21–9.18	2.68	62.7
Diallo ¹⁰	38	46	212	873	3.40	2.16–5.36	2.32	56.8
Simonsen ¹³	17	70	21	232	2.68	1.34–5.37	2.35	57.5
Tyndall ¹⁷	85	93	105	527	4.59	3.20–6.58	2.87	65.2
Nasio ¹⁶	86	78	137	580	4.67	3.26–6.68	2.74	63.6
Mehenda ¹²⁷	837	3411	38	253	1.63	1.15–2.32	1.51	33.7
Bollinger ²⁸	50	241	1	14	2.90	0.37–22.60	2.58	61.2
Chiasson ²²	36	797	14	542	1.75	0.93–3.27	1.72	41.7
TB patients								
Sassan ³⁰	75	18	415	221	2.22	1.29–3.81	1.24	19.1
High-risk total	1849	5207	1431	4751	1.18	1.09–1.28	1.13	11.7
Partner studies								
Hunter ³¹	43	330	165	3765	2.97	2.09–4.24	2.75	63.6
Carael ³³	90	105	34	45	1.13	0.67–1.92	1.07	6.8
Chao ³⁴	442	4844	75	232	0.28	0.21–0.37	0.34	65.8 [†]
Moss ³²	24	16	12	17	2.13	0.80–5.62	1.45	31.0
Allen ³⁵	275	612	132	324	1.10	0.86–1.41	1.07	6.6
Sedlin ³⁶	32	26	33	40	1.49	0.75–2.98	1.22	18.1
Konde-Luc ⁴⁵	153	1516	6	127	2.13	0.93–4.93	2.03	50.8
Partner total	1059	7449	457	4550	1.42	1.26–1.59	1.36	26.7
Random population surveys								
Barongo ³⁷	55	1356	42	642	0.62	0.41–0.94	0.63	36.5 [†]
Grosskurth ³⁸	158	4604	61	1026	0.58	0.43–0.78	0.59	40.9 [†]
Van de Perre ⁴⁰	46	224	6	26	0.89	0.35–2.28	0.91	9.1 [†]
Seed ⁴³	171	422	51	192	1.52	1.07–2.17	1.37	27.2
Malamba ⁴⁴	111	114	21	47	2.17	1.22–3.88	1.60	56.8
Quigley ⁴⁶	101	272	48	121	0.94	0.62–1.40	0.95	4.7 [†]
Urassa 1 ²⁶	56	1301	42	600	0.61	0.41–0.93	0.63	36.9 [†]
Urassa 2	105	2040	32	426	0.69	0.46–1.03	0.70	29.9 [†]
Urassa 3	38	309	19	158	1.02	0.57–1.83	1.02	2.0
Urassa 4	112	716	54	692	2.00	1.43–2.82	1.87	46.5
Urassa 5	101	365	48	136	0.78	0.53–1.16	0.83	16.9 [†]
Random total	1054	11723	424	4066	0.86	0.77–0.97	0.87	12.6 [†]
Total	3962	24379	2312	13367	0.94	0.89–0.99	0.95	5.2 [†]

OR = 1.06

Circumcision
increases risk of
HIV!

Howe 1999

Intact+=non-circumcised and HIV+. Intact- =non-circumcised and HIV- . Circ+=circumcised and HIV+. Circ- =circumcised and HIV- .

Case study 3. Does male circumcision reduce risk of HIV?

Circumcision in men and the prevention of HIV infection: a 'meta-analysis' revisited

Nigel O'Farrell MD MRCP¹ and Matthias Egger MD MFPHM²

¹*Department of Genitourinary Medicine, Milne Clinic, Bristol Royal Infirmary, Bristol, and*

²*MRC Health Services Research Collaboration, Department of Social Medicine, University of Bristol, Bristol, UK*

Summary: There is debate on the role of male circumcision in HIV transmission. Most case-control and cohort studies from Africa have shown an association between a lack of circumcision and an increased risk of HIV infection in men. The evidence is conflicting, however, with cross-sectional surveys from Tanzania and Rwanda either showing no relationship or an association in the opposite direction. A recent review and meta-analysis of the literature¹ concluded that the risk of HIV infection was lower in uncircumcised men (combined odds ratio 0.94, 95% confidence interval 0.89 to 0.99). However, the analysis was performed by simply pooling the data from 33 diverse studies, which is an inappropriate method for combining studies.

We re-analysed the data, stratifying by study, and found that an intact foreskin was associated with an increased risk of HIV infection: combined odds ratio 1.43 (1.32 to 1.54) with a fixed effect model and 1.67 (1.25 to 2.24) with a random effect model. There was significant between-study heterogeneity ($P < 0.0001$) which was partly explained by stronger associations in studies in high-risk groups. The results from this re-analysis thus support the contention that male circumcision may offer protection against HIV infection, particularly in high-risk groups where genital ulcers and other STDs 'drive' the HIV epidemic. A systematic review is required to clarify this issue. Such a review should be based on an extensive search for relevant studies, published and unpublished, and should include a careful assessment of the design and methodological quality of studies. Much emphasis should be given to the exploration of possible sources of heterogeneity. In view of the continued high prevalence and incidence of HIV in many countries in sub-Saharan Africa, the question of whether circumcision could contribute to prevent infections is of great importance, and a sound systematic review of the available evidence should be performed without delay.

Circumcision
lowers risk of
HIV!

Cochrane Review: Confounding is a major concern

HIV and male circumcision—a systematic review with assessment of the quality of studies

N Siegfried, M Muller, J Deeks, J Volmink, M Egger, N Low, S Walker, and P Williamson

This Cochrane systematic review assesses the evidence for an interventional effect of male circumcision in preventing acquisition of HIV-1 and HIV-2 by men through heterosexual intercourse. The review includes a comprehensive assessment of the quality of all 37 included observational studies. Studies in high-risk populations consisted of four cohort studies, 12 cross-sectional studies, and three case-control studies; general population studies consisted of one cohort study, 16 cross-sectional studies, and one case-control study. There is evidence of methodological heterogeneity between studies, and statistical heterogeneity was highly significant for both general population cross-sectional studies ($\chi^2=132.34$; degrees of freedom [df]=15; $p<0.00001$) and high-risk cross-sectional studies ($\chi^2=29.70$; df=10; $p=0.001$). Study quality was very variable and no studies measured the same set of potential confounding variables. Therefore, conducting a meta-analysis was inappropriate. Detailed quality assessment of observational studies can provide a useful visual aid to interpreting findings. Although most studies show an association between male circumcision and prevention of HIV, these results may be limited by confounding, which is unlikely to be adjusted for.

*Lancet Infect Dis 2005;
5: 165–73*

NS and JV are at the South African Cochrane Centre, Medical Research Council, South Africa; NS is currently a Nuffield Medical Fellow at The University of Oxford, Oxford, UK; JV is also at the Primary Health Care Directorate, University of Cape Town, Cape Town, South Africa; MM is at the Institute for Maritime Technology, Simon's Town, South Africa; JD is at the Centre for Statistics in Medicine, Institute of Health Sciences,

Finally, RCTs get published!

Open access, freely available online PLOS MEDICINE

Randomized, Controlled Intervention Trial of Male Circumcision for Reduction of HIV Infection Risk: The ANRS 1265 Trial

Bertran Auvert^{1,2,3,4*}, Dirk Taljaard⁵, Emmanuel Lagarde^{2,4}, Joëlle Sobngwi-Tambekou⁶, Rémi Sitta^{2,4}, Adrian Puren⁶

1 Hôpital Ambroise-Paré, Assistance Publique—Hôpitaux de Paris, Boulogne, France, **2** INSERM U 687, Saint-Maurice, France, **3** University Versailles Saint-Quentin, Versailles, France, **4** IIR 69, Villejuif, France, **5** Progressus, Johannesburg, South Africa, **6** National Institute for Communicable Disease, Johannesburg, South Africa

Competing Interests: The authors have declared that no competing interests exist.

Author Contributions: BA designed the study with DT, EJ, and AP. DT and AP were responsible for operational aspects, including laboratory and field work and in-country administration of the study. BA monitored the study with input from EJ, and wrote the paper with input from all authors. BA analyzed the data with HS, with inputs from JZT. RS conducted the interim analysis.

Academic Editor: Steven Derris, San Francisco General Hospital, San Francisco, California, United States of America.

Citation: Auvert B, Taljaard D, Lagarde E, Sobngwi-Tambekou J, Sitta R, et al. (2005) Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 Trial. *PLOS Med* 2(11): e298.

Received: June 29, 2005

Accepted: September 26, 2005

Published: October 25, 2005

DOI:

10.1371/journal.pmed.0020298

Copyright: © 2005 Auvert et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background

Observational studies suggest that male circumcision may provide protection against HIV-1 infection. A randomized, controlled intervention trial was conducted in a general population of South Africa to test this hypothesis.

Methods and Findings

A total of 3,274 uncircumcised men, aged 18–24 y, were randomized to a control or an intervention group with follow-up visits at months 3, 12, and 21. Male circumcision was offered to the intervention group immediately after randomization and to the control group at the end of the follow-up. The grouped censored data were analyzed in intention-to-treat, univariate and multivariate, analyses, using piecewise exponential, proportional hazards models. Rate ratios (RR) of HIV incidence were determined with 95% CI. Protection against HIV infection was calculated as 1 – RR. The trial was stopped at the interim analysis, and the mean interquartile range follow-up was 18.1 mo (13.0–21.0) when the data were analyzed. There were 20 HIV infections (incidence rate = 0.85 per 100 person-years) in the intervention group and 49 (2.1 per 100 person-years) in the control group, corresponding to an RR of 0.40 (95% CI: 0.24–0.68; $p < 0.001$). This RR corresponds to a protection of 60% (95% CI: 32%–76%). When controlling for behavioural factors, including sexual behaviour that increased slightly in the intervention group, condom use, and health-seeking behaviour, the protection was of 61% (95% CI: 34%–77%).

Conclusion

Male circumcision provides a degree of protection against acquiring HIV infection, equivalent to what a vaccine of high efficacy would have achieved. Male circumcision may provide an important way of reducing the spread of HIV infection in sub-Saharan Africa. (Preliminary and partial results were presented at the International AIDS Society 2005 Conference, on 26 July 2005, in Rio de Janeiro, Brazil.)

Male circumcision for HIV prevention in young men in Kisumu, Kenya: a randomised controlled trial

Robert C Bailey, Stephen Moses, Corette B Parker, Ian Maclean, John N Krieger, Carolyn F M Williams, Richard T Campbell, Jekoonah O Ndirya-Achola

Summary

Background Male circumcision could provide substantial protection against acquisition of HIV-1 infection. Our aim was to determine whether male circumcision had a protective effect against HIV infection, and to assess safety and changes in sexual behaviour related to this intervention.

Methods We did a randomised controlled trial of 2784 men aged 18–24 years in Kisumu, Kenya. Men were randomly assigned to an intervention group (circumcision; $n=1391$) or a control group (delayed circumcision, 1393), and assessed by HIV testing, medical examinations, and behavioural interviews during follow-ups at 1, 3, 6, 12, 18, and 24 months. HIV seroconversion was estimated in an intention-to-treat analysis. This trial is registered with ClinicalTrials.gov, with the number NCT00059371.

Findings The trial was stopped early on December 12, 2006, after a third interim analysis reviewed by the data and safety monitoring board. The median length of follow-up was 24 months. Follow-up for HIV status was incomplete for 240 (8.6%) participants. 22 men in the intervention group and 47 in the control group had tested positive for HIV when the study was stopped. The 2-year HIV incidence was 2.1% (95% CI 1.2–3.0) in the circumcision group and 4.2% (3.0–5.4) in the control group ($p=0.0065$); the relative risk of HIV infection in circumcised men was 0.47 (0.28–0.78), which corresponds to a reduction in the risk of acquiring an HIV infection of 53% (22–72). Adjusting for non-adherence to treatment and excluding four men found to be seropositive at enrolment, the protective effect of circumcision was 60% (32–77). Adverse events related to the intervention (21 events in 1.5% of those circumcised) resolved quickly. No behavioural risk compensation after circumcision was observed.

Interpretation Male circumcision significantly reduces the risk of HIV acquisition in young men in Africa. Where appropriate, voluntary, safe, and affordable circumcision services should be integrated with other HIV preventive interventions and provided as expeditiously as possible.

Male circumcision for HIV prevention in men in Rakai, Uganda: a randomised trial

Ronald H Gray, Geoffrey Kigazi, David Serwadda, Frederick Makumbi, Stephen Watya, Fred Nalugoda, Noah Kiwanuka, Lawrence H Moulton, Mohammad A Choudhary, Michael Z Chen, Nelson K Sewankambo, Fred Wabwire-Mangen, Melanie C Bacon, Carolyn F M Williams, Pius Opendi, Steven J Reynolds, Oliver Laeyendecker, Thomas C Quinn, Maria J Wawer

Summary

Background Ecological and observational studies suggest that male circumcision reduces the risk of HIV acquisition in men. Our aim was to investigate the effect of male circumcision on HIV incidence in men.

Methods 4996 uncircumcised, HIV-negative men aged 15–49 years who agreed to HIV testing and counselling were enrolled in this randomised trial in rural Rakai district, Uganda. Men were randomly assigned to receive immediate circumcision ($n=2474$) or circumcision delayed for 24 months (2522). HIV testing, physical examination, and interviews were repeated at 6, 12, and 24 month follow-up visits. The primary outcome was HIV incidence. Analyses were done on a modified intention-to-treat basis. This trial is registered with ClinicalTrials.gov, with the number NCT00425984.

Findings Baseline characteristics of the men in the intervention and control groups were much the same at enrolment. Retention rates were much the same in the two groups, with 90–92% of participants retained at all time points. In the modified intention-to-treat analysis, HIV incidence over 24 months was 0.66 cases per 100 person-years in the intervention group and 1.33 cases per 100 person-years in the control group (estimated efficacy of intervention 51%, 95% CI 16–72; $p=0.006$). The as-treated efficacy was 55% (95% CI 22–75; $p=0.002$); efficacy from the Kaplan-Meier time-to-HIV-detection as-treated analysis was 60% (30–77; $p=0.003$). HIV incidence was lower in the intervention group than it was in the control group in all sociodemographic, behavioural, and sexually transmitted disease symptom subgroups. Moderate or severe adverse events occurred in 84 (3.6%) circumcisions; all resolved with treatment. Behaviours were much the same in both groups during follow-up.

Interpretation Male circumcision reduced HIV incidence in men without behavioural disinhibition. Circumcision can be recommended for HIV prevention in men.



Lenet 2007; 369: 657–66

See Editorial page 615

See Comment page 627

See Perspectives page 635

See Articles page 643

See Viewpoint page 708

Johns Hopkins University,

Bloomberg School of Public

Health, Baltimore, MD, USA

(Prof R H Gray MD,

Prof L H Moulton PhD,

M A Chaudhary PhD,

M Z Chen MSc,

Prof M J Wawer MD's Rakai

Health Sciences Program,

Entebbe, Uganda

(C Kigazi MSc),

F Nalugoda MHS,

N Kiwanuka MSc),

P Opendi MSc), Makerere

University, Institute of Public

Health, Kampala, Uganda

(D Serwadda MSc),

F Wabwire-Mangen PhD's,

Makere University, Mulago



Lenet 2007; 369: 643–66

See Editorial page 615

See Comment page 617

See Articles page 657

See Viewpoint page 708

Division of Epidemiology and

Biostatistics, University of

Illinois at Chicago, Chicago, IL,

USA (Prof R C Bailey PhD,

Prof R T Campbell PhD),

Department of Medical

Microbiology (I Maclean PhD),

Community Health Sciences

and Medicine

(Prof S Moses MD), UNM

Project, Kisumu, Kenya and

Department of Community

Health Sciences (K Agot PhD),

University of Manitoba,

Winnipeg, Canada; RTI

International, Research

Triangle Park, NC, USA

(C B Parker DPhil), Department

of Urology, University of

Washington School of

Medicine, Seattle, WA, USA

(Prof N Krieger MD), Division

Meta-analysis of 3 RCTs shows strong, consistent effect

Table 1 Study characteristics and outcomes

Study	Design	Setting	Population	n	Outcomes		Relative risk (95% confidence interval)
					Intervention	Control	
Auvert <i>et al.</i> (2005)*	Randomized trial	Orange farm, South Africa	Males aged 18–24 years	3128	20/1546	49/1582	0.42 (0.25–0.70)
Bailey <i>et al.</i> (2007)**	Randomized trial	Kisumu, Kenya	Males aged 18–24 years	2780	19/1388	46/1392	0.41 (0.24–0.70)
Gray <i>et al.</i> (2007)	Randomized trial	Rakai district, Uganda	Males aged 15–49 years	4996	22/2474	45/2522	0.50 (0.30–0.83)

*3274 randomized, 3128 included in analysis.

**2784 randomized, 2780 included in analysis.

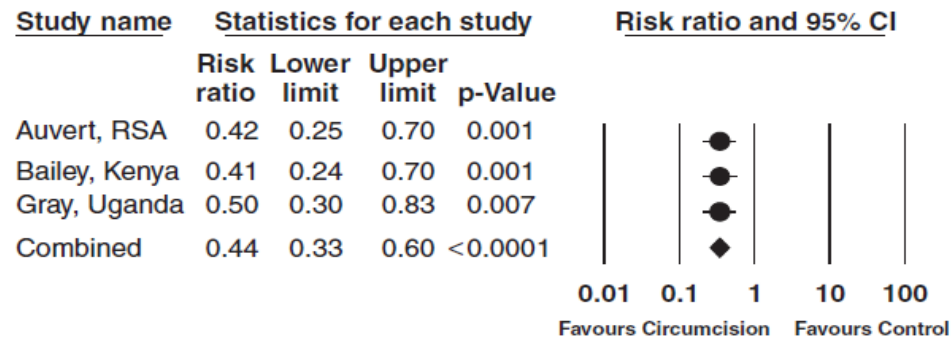


Fig. 1 Random effects meta-analysis.

What is evidence-based medicine?

The practice of EBM is the integration of

- individual clinical expertise
with the
- best available external clinical evidence
from systematic research and
- patient's values and expectations

The importance of research synthesis

- ☞ We need evidence for both clinical practice and for public health decision making
- ☞ Where does evidence come from?
 - An good review is a state-of-the-art synthesis of current evidence on a given research question
 - Given the explosion of medical literature, and the fact that time is always scarce, review articles play a big role in decision-making
 - According to one estimate, to keep up to date in Internal Medicine, need to read 17 articles a day, 365 days a year!

The importance of research synthesis

- Given that most clinicians and public health professionals do not have the time to track down all the original articles, critically read them, and obtain the evidence they need for their questions,
 - Systematic reviews and clinical practice guidelines may be their best source of evidence
 - Several “pre-digested” sources of evidence are currently available
 - The EBM movement is heavily dependent on these pre-appraised evidence sources

Hierarchy of evidence

Grade of Recommendation	Level of Evidence	Therapy/Prevention, Aetiology/Harm	Prognosis	Diagnosis	Economic analysis
A	1a	SR (with homogeneity*) of RCTs	SR (with homogeneity*) of inception cohort studies; or a CPG [†] validated on a test set.	SR (with homogeneity*) of Level 1 diagnostic studies; or a CPG validated on a test set.	SR (with homogeneity*) of Level 1 economic studies
	1b	Individual RCT (with narrow Confidence Interval [‡])	Individual inception cohort study with $\geq 80\%$ follow-up	Independent blind comparison of an appropriate spectrum of consecutive patients, all of whom have undergone both the diagnostic test and the reference standard.	Analysis comparing all (critically-validated) alternative outcomes against appropriate cost measurement, and including a sensitivity analysis incorporating clinically sensible variations in important variables.
	1c	All or none [§]	All or none case-series [™]	Absolute SpPins and SnNouts ^{††}	Clearly as good or better ^{‡‡} , but cheaper. Clearly as bad or worse but more expensive. Clearly better or worse at the same cost.
B	2a	SR (with homogeneity*) of cohort studies	SR (with homogeneity*) of either retrospective cohort studies or untreated control groups in RCTs.	SR (with homogeneity*) of Level ≥ 2 diagnostic studies	SR (with homogeneity*) of Level ≥ 2 economic studies
	2b	Individual cohort study (including low quality RCT; e.g., $<80\%$ follow-up)	Retrospective cohort study or follow-up of untreated control patients in an RCT; or CPG not validated in a test set.	Independent blind comparison but either in non-consecutive patients, or confined to a narrow spectrum of study individuals (or both), all of whom have undergone both the diagnostic test and the reference standard; or a diagnostic CPG not validated in a test set.	Analysis comparing a limited number of alternative outcomes against appropriate cost measurement, and including a sensitivity analysis incorporating clinically sensible variations in important variables.
	2c	"Outcomes" Research	"Outcomes" Research		
	3a	SR (with homogeneity*) of case-control studies			
	3b	Individual Case-Control Study		Independent blind comparison of an appropriate spectrum, but the reference standard was not applied to all study patients	Analysis without accurate cost measurement, but including a sensitivity analysis incorporating clinically sensible variations in important variables.
C	4	Case-series (and poor quality cohort and case-control studies ^{§§})	Case-series (and poor quality prognostic cohort studies ^{™™})	Reference standard was not applied independently or not applied blindly	Analysis with no sensitivity analysis
D	5	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on physiology, bench research or "first principles"	Expert opinion without explicit critical appraisal, or based on economic theory

Guidelines and recommendations: GRADE

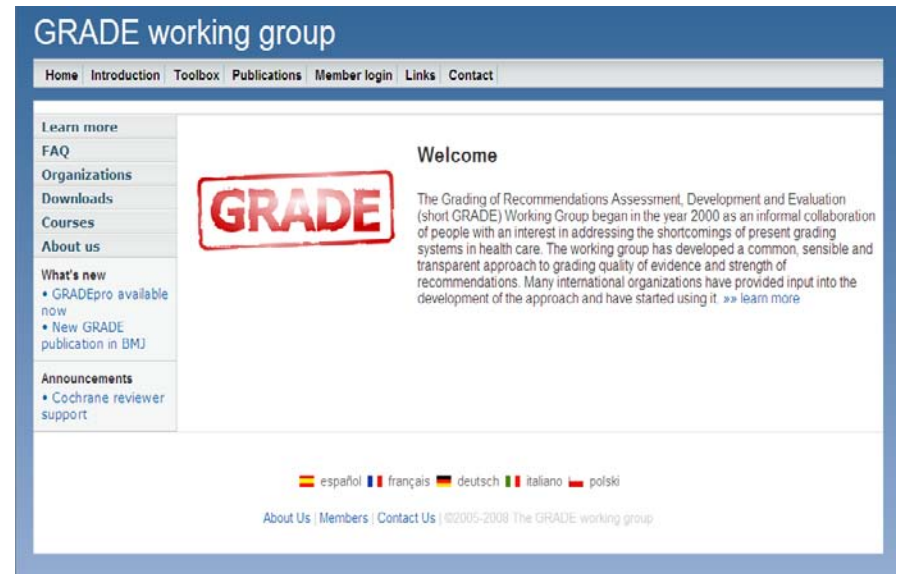
ANALYSIS

Downloaded from bmj.com on 18 May 2008

RATING QUALITY OF EVIDENCE AND STRENGTH OF RECOMMENDATIONS

GRADE: an emerging consensus on rating quality of evidence and strength of recommendations

Guidelines are inconsistent in how they rate the quality of evidence and the strength of recommendations. This article explores the advantages of the GRADE system, which is increasingly being adopted by organisations worldwide



Doing New Research? Don't Forget the Old

Nobody should do a trial without reviewing what is known

Mike Clarke

On May 2, 1898, George Gould used his address to the founding meeting of the Association of Medical Librarians in Philadelphia to present a vision of the future of health information. 'I look forward,' he said, 'to such an organisation of the literary records of medicine that a puzzled worker in any part of the civilised world shall in an hour be able to gain a knowledge pertaining to a subject of the experience of every other man in the world' [1]. Has his vision been realised?

good quality, but some of it is not. Thus, anyone wishing to use the health literature to make well-informed decisions must both identify the relevant research from amidst this vast amount of information and then appraise it. This is an impossible task for many. Even though making access to the literature easier and cheaper will increase the ability of people to find research, it will also reveal just how much information there is out there and how daunting is the task of making sense of it.

with one or more search engines? Almost certainly, as the speed of the search increased through these four

Citation: Clarke M (2004) Doing new research? Don't forget the old. PLoS Med 1(2):e35.

Copyright: © 2004 access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, any medium, provided the original author and source are credited.

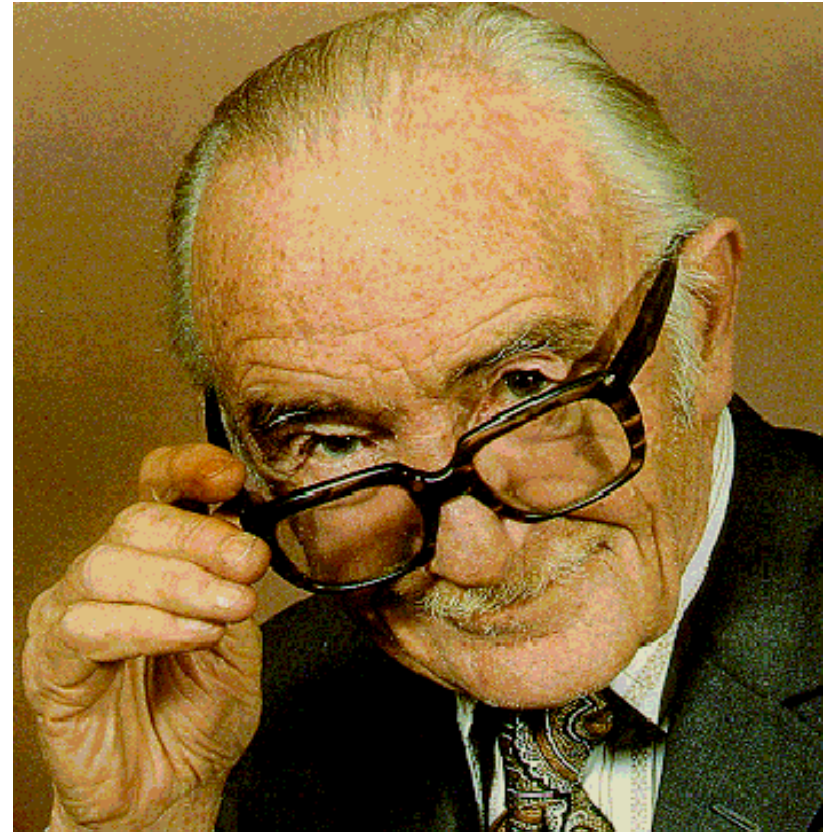
Mike Clarke is director, Cochrane Centre, mclarke@cochrane.org

Box 1. Practical Suggestions for Researchers

- Conduct a systematic review of your research question before embarking on a new study, or identify a relevant review done by someone else.
- Design your study to take account of the relevant successes and failures of the prior studies, and of the evidence within them.
- Discuss the findings of your study in the context of an updated systematic review of relevant research.
- Publish the systematic review within, alongside, or shortly after the report of your study.
- Provide information from your study to others doing systematic reviews of similar topics.

Prof Archibald Cochrane, CBE (1909 - 1988)

- ✧ The Cochrane Collaboration is named in honour of Archie Cochrane, a British researcher.
- ✧ In 1979 he wrote, *"It is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials"*



The Cochrane Collaboration

- Archie Cochrane's challenge led to the establishment during the 1980s of an international collaboration to develop the *Oxford Database of Perinatal Trials*.
- His encouragement, and the endorsement of his views by others, led to the opening of the first Cochrane centre (in Oxford, UK) in 1992 and the founding of The **Cochrane Collaboration** in 1993.





The Cochrane Library

Evidence for healthcare decision-making

BROWSE

Cochrane Reviews: [By Topic](#) | [New Reviews](#) | [Updated Reviews](#) | [A-Z](#) | [By Review Group](#)
Other Resources: [Other Reviews](#) | [Clinical Trials](#) | [Methods Studies](#) | [Technology Assessments](#) | [Economic Evaluations](#)

[? More Info](#)

SEARCH

 [Advanced Search](#) | [MeSH Search](#) | [Search History](#) | [Saved Searches](#)

Welcome to The Cochrane Library

The Cochrane Library contains high-quality, independent evidence to inform healthcare decision-making. It includes reliable evidence from Cochrane and other systematic reviews, clinical trials, and more. Cochrane reviews bring you the combined results of the world's best medical research studies, and are recognised as the gold standard in evidence-based health care.

More About The Cochrane Library

[What are Systematic Reviews & Protocols?](#) | [Product Descriptions](#) | [About the CD-ROM](#)

What's New in Issue 2, 2008?

[Important changes to The Cochrane Library \(PDF\)](#)[Highlights of new and updated Reviews \(PDF\)](#)[Most Accessed Reviews 2007 \(PDF\)](#)[Antioxidant supplements for prevention of mortality in healthy participants and patients with various diseases](#)[Anti-histamines for prolonged non-specific cough in children](#)[Exercise for the management of cancer-related fatigue in adults](#)[Effect of timing of umbilical cord clamping of term infants on maternal and neonatal outcomes](#)[Dopamine agonist therapy in early Parkinson's disease](#)[House dust mite control measures for asthma](#)[Interventions for enhancing medication adherence](#)[Cognitive-behavioural interventions for preventing youth gang involvement for children and young people \(7-16\)](#)[Interventions for treating wrist fractures in children](#)[Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment](#)[Short-course versus long-course antibiotic therapy for non-severe community-acquired pneumonia in children aged 2 months to 59 months](#)[Protocols for Cochrane Reviews of Diagnostic Test Accuracy](#)[New Reviews](#) | [Updated Reviews](#) | [Press Room](#) | [Podcasts](#)

Access to The Cochrane Library

The Cochrane Library is available online through Wiley InterScience.

[More About Access to Cochrane](#)

Help! New Users Start Here

As a new user we recommend you use the following resources to help you navigate through the evidence and get the most out of The Cochrane Library. [More](#)

For Clinicians

As a clinician you are under constant pressure to have high-quality, up-to-date evidence at your fingertips. [More](#)

For Researchers

The internet has given us instant access to a huge amount of research, but the large volume of available information is a problem in itself. [More](#)

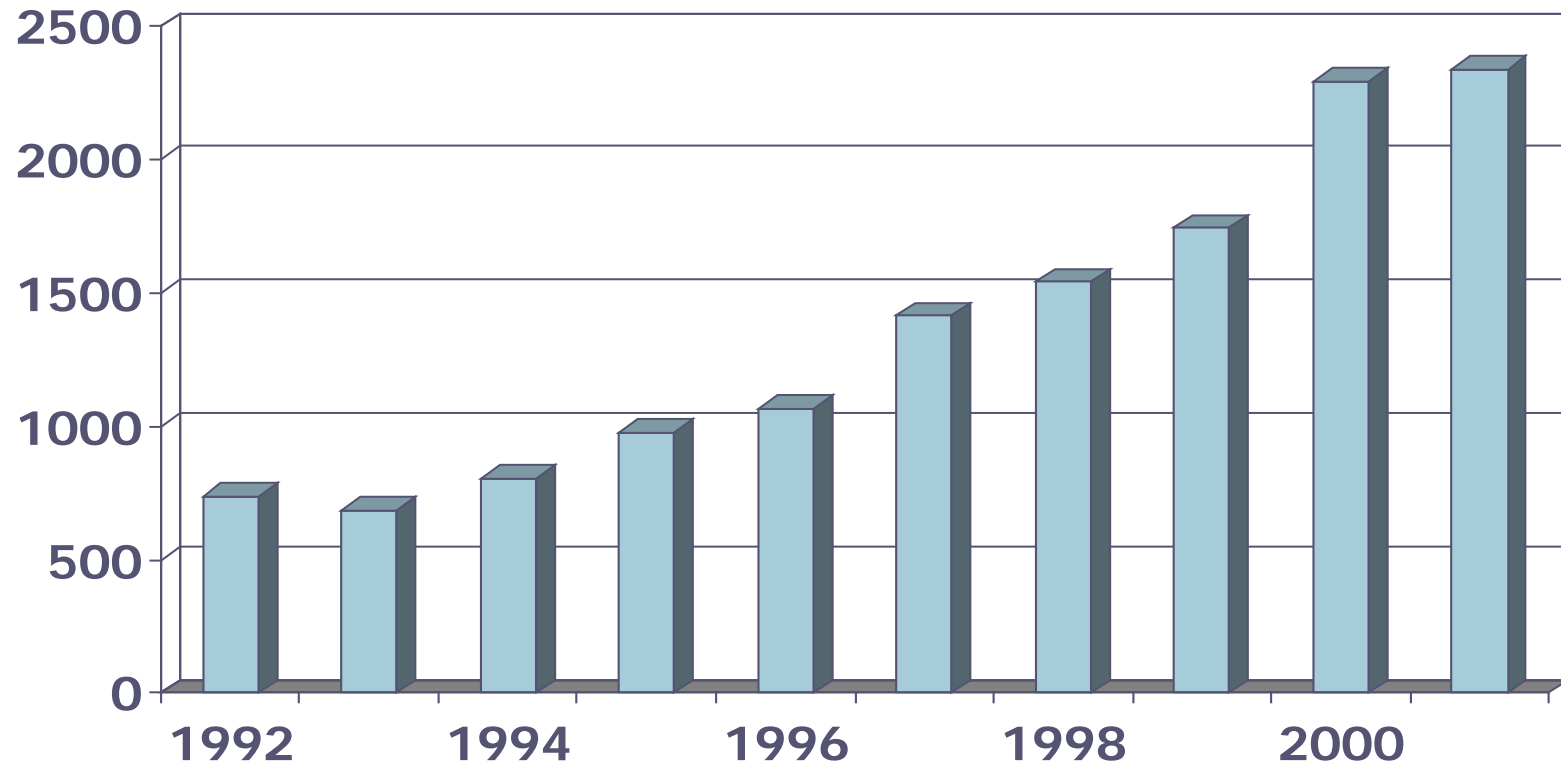
For Patients

Healthcare consumers and patients need high-quality evidence about the effectiveness of treatments. [More](#)

For Policy Makers

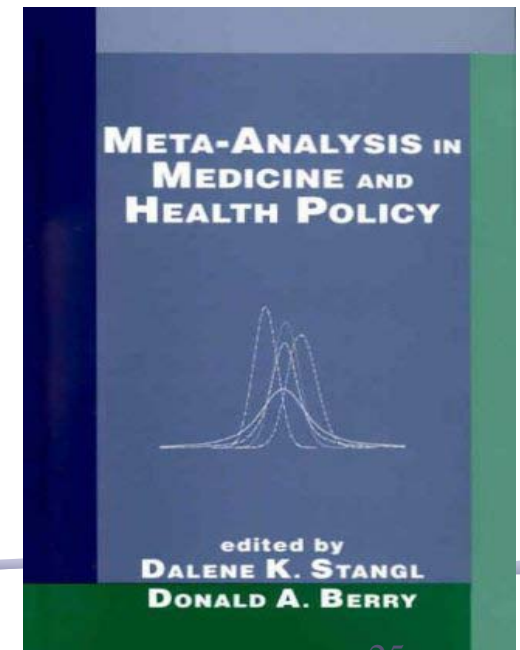
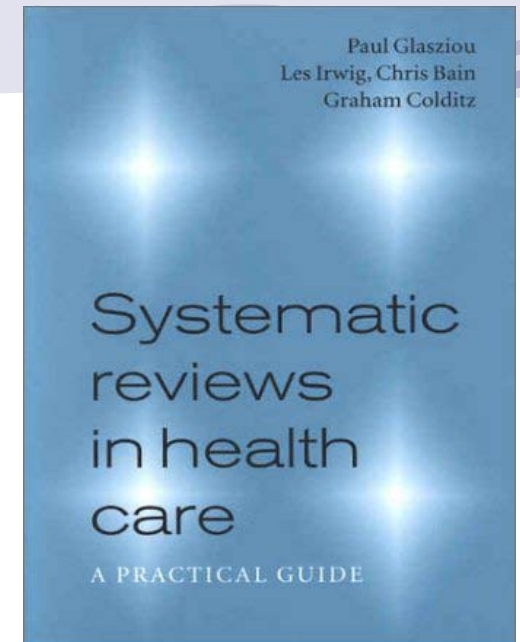
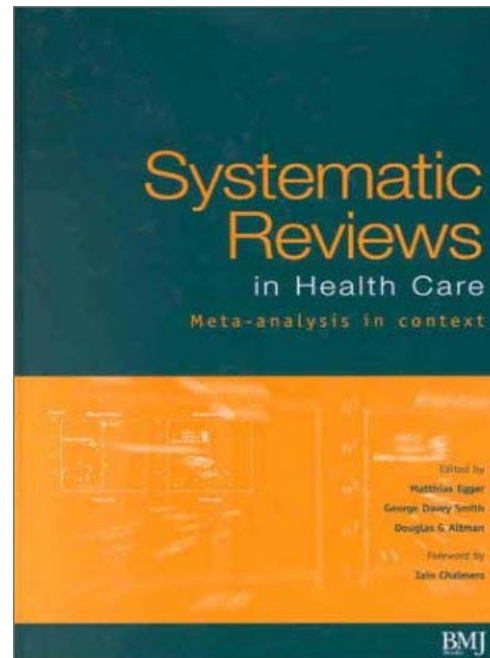
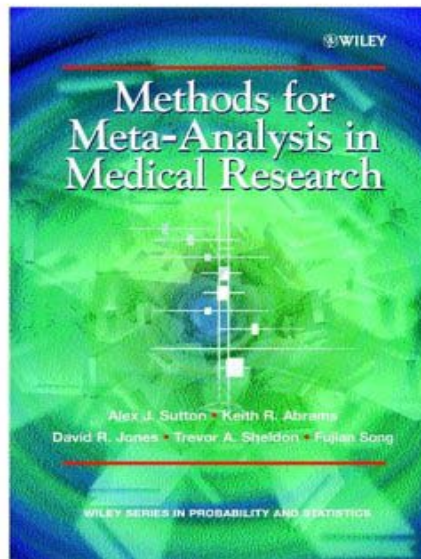
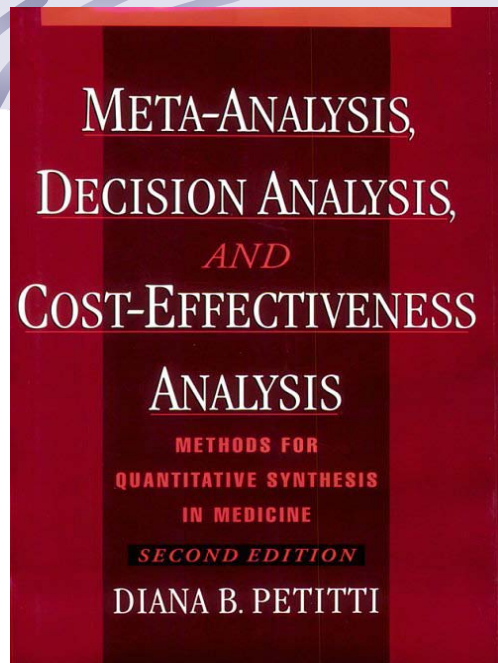
As a policy maker or healthcare manager you are a generalist in search of high-quality information across a broad range of issues. [More](#)

Systematic reviews/meta-analyses indexed in PubMed – 10 years



Search: meta-analysis(MeSH) OR meta-analysis(tw) OR systematic review(tw)

2500 SRs per year, of which about 20% are Cochrane reviews (estimate by Moher et al. PLoS Med 2007)



ACP Journal Club

March/April 1997
Volume 126 • Number 2

Linking Research to Practice in Internal Medicine

Published
Bimonthly
by the
American College
of Physicians

■ THERAPEUTICS

Lowering LDL cholesterol levels reduced fatal coronary events in patients with acute MI and average cholesterol levels	29
Amlodipine did not increase morbidity or mortality rates in severe heart failure	30
Meta-analysis: Mortality is reduced when fibrinolytic therapy is started soon after the onset of MI symptoms	31
Hirudin was no more effective than unfractionated heparin for acute MI	32
Hirudin reduced death or MI more than heparin at 48 hours but not at 30 days	33
Meta-analysis: β -blockers improve function in dilated cardiomyopathy	34
Meta-analysis: Thrombolytic therapy increases the risk for early death and intracranial hemorrhage after acute ischemic stroke	35
Meta-analysis: Misoprostol reduces NSAID-induced gastrointestinal mucosal injury	36
Meta-analysis: Pentoxifylline improves walking in intermittent claudication	37
Meta-analysis: Respiratory rehabilitation relieves dyspnea in COPD	38
Review: Antibiotics are ineffective for acute bronchitis	39
Implantable insulin pump improved quality of life in NIDDM	40
Valacyclovir accelerated healing in recurrent herpes genitalis	41
Subcutaneous tunneling reduced catheter-related sepsis in critically ill patients	42

■ DIAGNOSIS

D-dimer levels detected DVT in patients hospitalized for stroke rehabilitation	43
CSF proteins 150 and 131 were specific for diagnosing Creutzfeldt-Jakob disease	44
Brain protein 14-3-3 was a sensitive test for Creutzfeldt-Jakob disease	45
Meta-analysis: Glycosylated hemoglobin levels are useful for diagnosing diabetes	46

■ PROGNOSIS

Ischemic stroke with accompanying atrial fibrillation was associated with reduced survival and functional status	47
------------------------------------------------------------------------------------------------------------------	----

■ ETIOLOGY

LDL particle size was smaller in CAD, but other lipid parameters were stronger predictors of CAD	48
Triglyceride level but not LDL particle size was an independent risk factor for MI in men	49
Low cholesterol levels were associated with suicide in men	50

■ ECONOMICS

Low-molecular-weight heparin was cost-effective for perioperative prevention of DVT	51
Cost-effectiveness of pravastatin in the secondary prevention of CAD in men varied with the risk profile of the patient	52
Intensive therapy extended life and was cost-effective for IDDM	53

Other Articles Noted	54
Glossary	56
Contents	A-7
Purpose and Procedure	A-13
Editorial: Transferring evidence from research into practice: 3. Developing evidence-based clinical policy	A-14

AJCEA 126(Na 2)29-56(1997)US ISSN 1056-8751

Vol. 2 No. 5 pp. 129-160

SEPTEMBER/OCTOBER 1997

Evidence-Based Medicine



American College
of Physicians

BMJ
Publishing
Group

Search

☒ All these words

☐ Any of these words





(Searches using AND/OR/NOT combinations override the above)

search

Year published From -

[restrictions info](#)

10 results per page

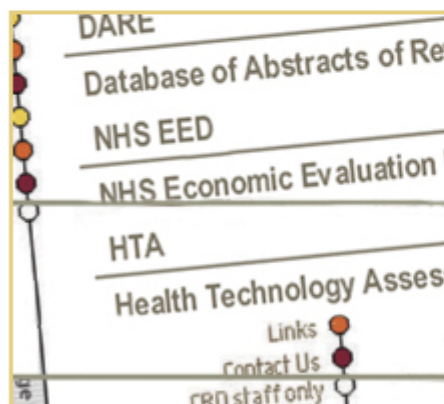
-  [Home](#)
-  [Results](#)
-  [MeSH](#)
-  [Search history](#)

All results

DARE

NHS EED

HTA



Centre for Reviews and Dissemination databases

CRD was established in January 1994, and produces and promotes the use of research based knowledge in health and social care.

DARE – (*Database of Abstracts of Reviews of Effects*) contains over 4000 abstracts of quality assessed and critically appraised systematic reviews. The database focuses on the effects of interventions used in health and social care.

NHS Economic Evaluation Database (NHS EED) contains over 6000 abstracts of quality assessed economic evaluations. The database aims to assist decision-makers by systematically identifying and describing economic evaluations, appraising their quality and highlighting their relative strengths and weaknesses.

Both *DARE* and *NHS EED* include details of abstracts in the process of being written and these can be 'fast-tracked' on request.

The *HTA* database brings together details of completed and ongoing health technology assessments from around the world. The abstracts in the database are descriptive rather than analytical and do not form critical appraisals of the reports. The database is produced in collaboration with the INAHTA Secretariat, based at SBU, Sweden.

Systematic reviews are done in different domains

OPEN ACCESS Freely available online

PLOS MEDICINE

Tuberculosis among Health-Care Workers in Low- and Middle-Income Countries: A Systematic Review

Rajnish Joshi^{1,2}, Arthur L. Reingold¹, Dick Menzies³, Madhukar Pai^{3*}

1 Division of Epidemiology, School of Public Health, University of California Berkeley, Berkeley, California, United States of America, 2 Department of Medicine, Mahatma Gandhi Institute of Medical Sciences, Sevagram, Maharashtra, India, 3 Montreal Chest Institute, McGill University, Montreal, Canada

Meta-analysis of “rates”

Systematic reviews are done in different domains

Annals of Internal Medicine

REVIEW

Systematic Review: T-Cell–based Assays for the Diagnosis of Latent Tuberculosis Infection: An Update

Madhukar Pai, MD, PhD; Alice Zwerling, MSc; and Dick Menzies, MD, MSc

Eur Respir J 2008; 32: 1165–1174
DOI: 10.1183/09031936.00061808
Copyright©ERS Journals Ltd 2008



GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis

D.I. Ling*, A.A. Zwerling* and M. Pai*,#

Meta-analysis of “diagnostic accuracy [diagnosis]”

Systematic reviews are done in different domains

Tropical Medicine and International Health

doi:10.1111/j.1365-3156.2006.01571.x

VOLUME 11 NO 6 PP 789–799 JUNE 2006

Chloroquine or amodiaquine combined with sulfadoxine–pyrimethamine for uncomplicated malaria: a systematic review

Jimee Hwang¹, Edward Bitarakwate², Madhukar Pai³, Arthur Reingold³, Philip J. Rosenthal⁴ and Grant Dorsey⁴

1 Department of Internal Medicine, University of California San Francisco, San Francisco, CA, USA

2 Elizabeth Glaser Pediatric AIDS Foundation, Kampala, Uganda

3 Division of Epidemiology, School of Public Health, University of California Berkeley, Berkeley, CA, USA

4 Department of Infectious Diseases, University of California San Francisco, San Francisco, CA, USA

OPEN ACCESS Freely available online

PLOS MEDICINE

Effect of Duration and Intermittency of Rifampin on Tuberculosis Treatment Outcomes: A Systematic Review and Meta-Analysis

Dick Menzies^{1*}, Andrea Benedetti¹, Anita Paydar¹, Ian Martin¹, Sarah Royce², Madhukar Pai¹, Andrew Vernon³, Christian Lienhardt⁴, William Burman⁵

1 Respiratory and Epidemiology Clinical Research Unit, Montreal Chest Institute & Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, Canada, 2 University of California at San Francisco, San Francisco, California, United States of America, 3 Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America, 4 International Union against Tuberculosis and Lung Diseases, and Institut de Recherche pour le Développement, Paris, France, 5 Denver Public Health, Denver, Colorado, United States of America

Meta-analysis of “RCTs [therapy]”

Systematic reviews are done in different domains

REVIEW ARTICLE

Risk of Tuberculosis From Exposure to Tobacco Smoke

A Systematic Review and Meta-analysis

Michael N. Bates, PhD; Asheena Khalakdina, PhD; Madhukar Pai, MD, PhD;
Lisa Chang, MPH; Fernanda Lessa, MD, MPH; Kirk R. Smith, PhD

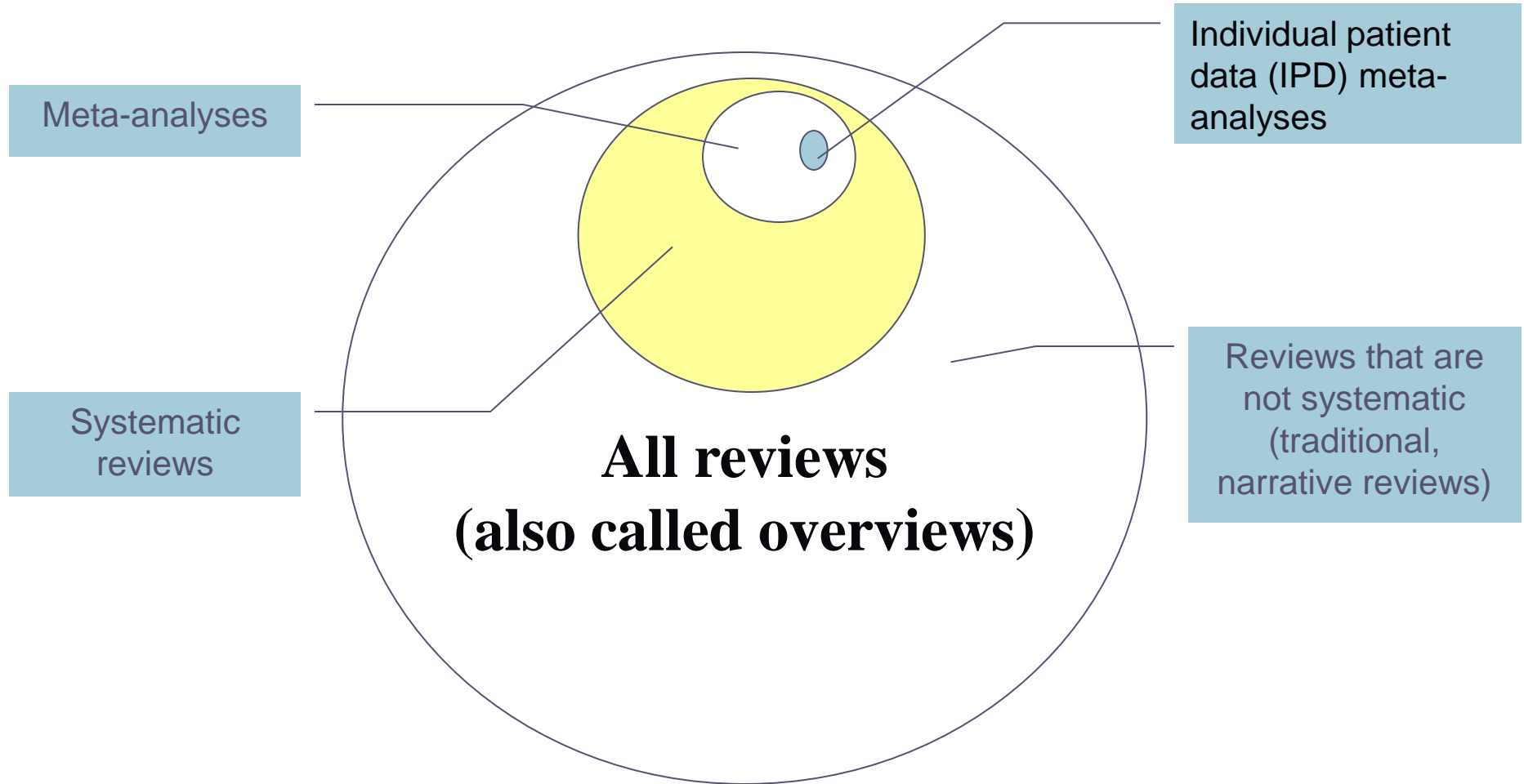
Arch Intern Med. 2007;167:335-342

Meta-analysis of “observational studies [etiology]”

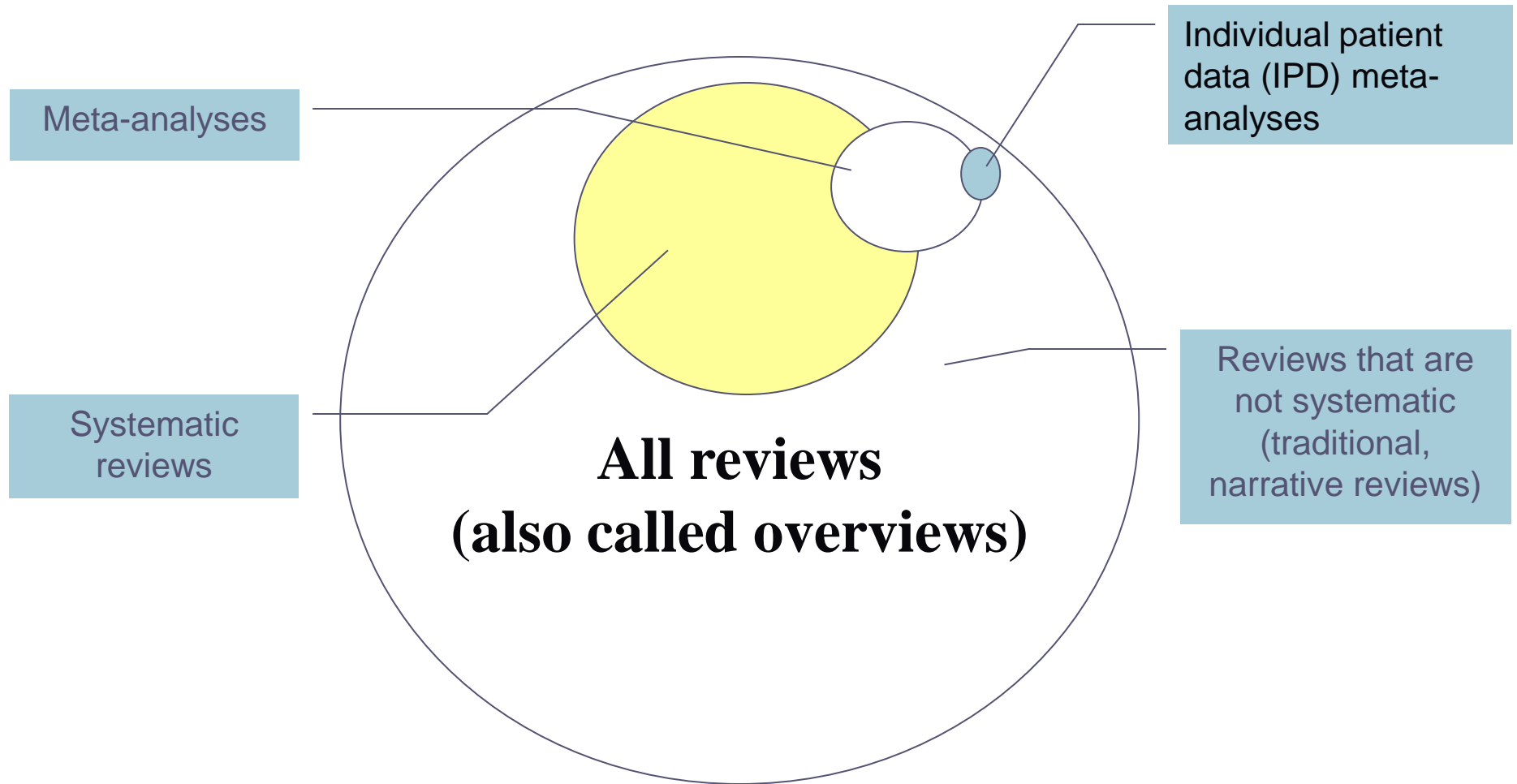
Are these the same or different?

- ☛ Traditional, narrative review
- ☛ Systematic review
- ☛ Overview
- ☛ Meta-analysis
- ☛ Pooled analysis

Types of review articles



In practice, not all meta-analyses are conducted as part of systematic reviews



Some definitions

- Traditional, narrative reviews, usually written by experts in the field, are qualitative, narrative summaries of evidence on a given topic. Typically, they involve informal and subjective methods to collect and interpret information.
- “A systematic review is a review in which there is a comprehensive search for relevant studies on a specific topic, and those identified are then appraised and synthesized according to a predetermined and explicit method.”*

Some definitions

- “A meta-analysis is the statistical combination of at least 2 studies to produce a single estimate of the effect of the healthcare intervention under consideration.” *
- Individual patient data meta-analyses (pooled analyses) involves obtaining raw data on all patients from each of the trials directly and then re-analyzing them.

Narrative vs. Systematic Reviews

TABLE I. Comparison of traditional and systematic reviews

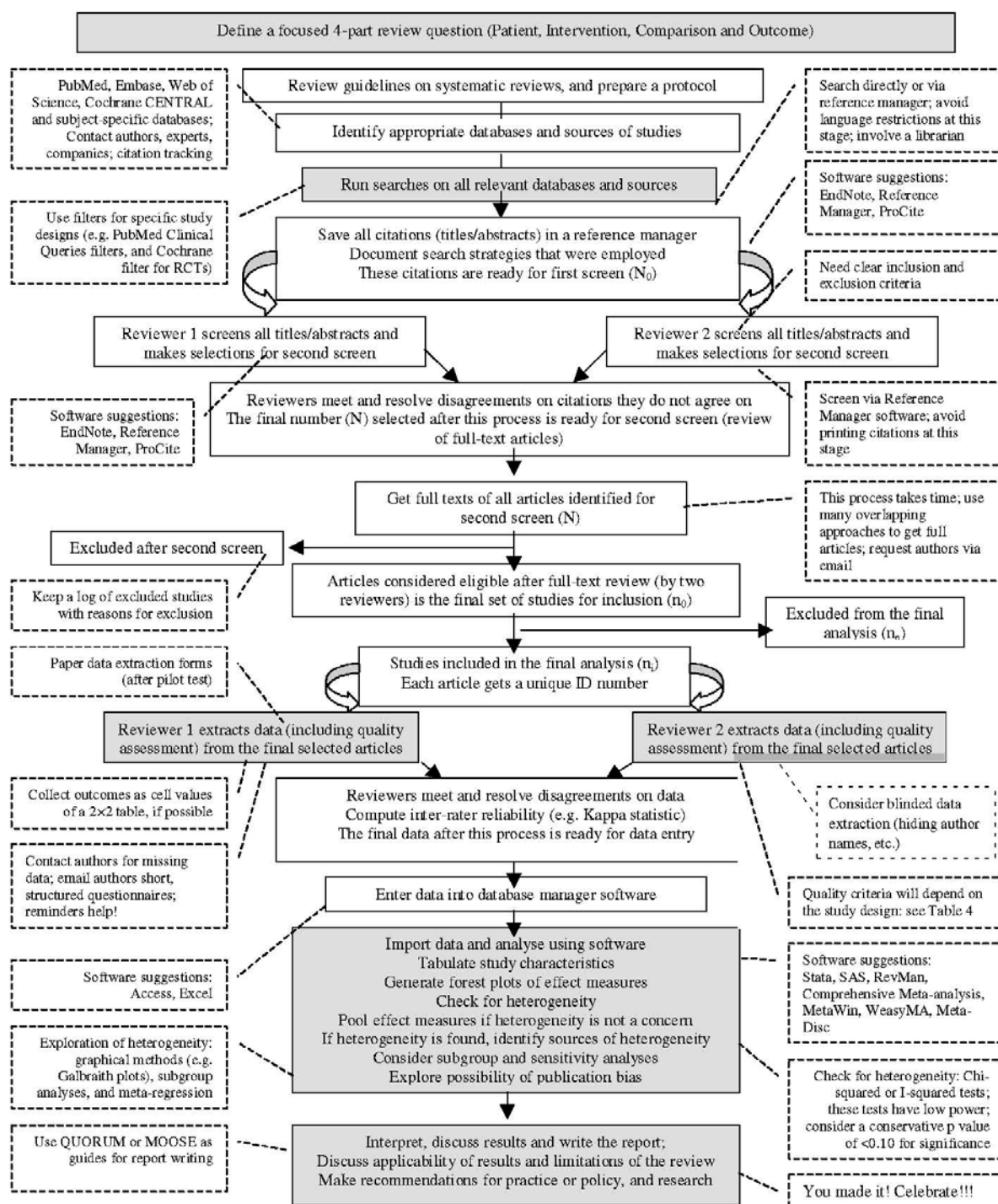
Components of a review	Traditional, narrative reviews	Systematic reviews
Formulation of the question	Usually address broad questions	Usually address focused questions
Methods section	Usually not present, or not well-described	Clearly described with pre-stated criteria about participants, interventions and outcomes
Search strategy to identify studies	Usually not described; mostly limited by reviewers' abilities to retrieve relevant studies; usually not reproducible and prone to selective citation	Clearly described and usually exhaustive; transparent, reproducible and less prone to selective citation
Quality assessment of identified studies	Usually all identified studies are included without explicit quality assessment	Only high-quality studies are included using pre-stated criteria; if lower-quality studies included, the effects of this are tested in subgroup analyses
Data extraction	Methods usually not described	Usually undertaken by more than one reviewer onto pre-tested data forms; attempts often made to obtain missing data from authors of primary studies
Data synthesis	Qualitative description employing the 'vote counting' approach, where each included study is given equal weight, irrespective of study size and quality	Meta-analysis assigns higher weights to effect measures from more precise studies; pooled, weighted effect measures with confidence limits provide power and precision to results
Heterogeneity	Usually dealt with in a narrative fashion	Heterogeneity dealt with by graphical and statistical methods; attempts are often made to identify sources of heterogeneity
Interpreting results	Prone to cumulative systematic biases and personal opinion	Less prone to systematic biases and personal opinion

Elements of a Systematic Review

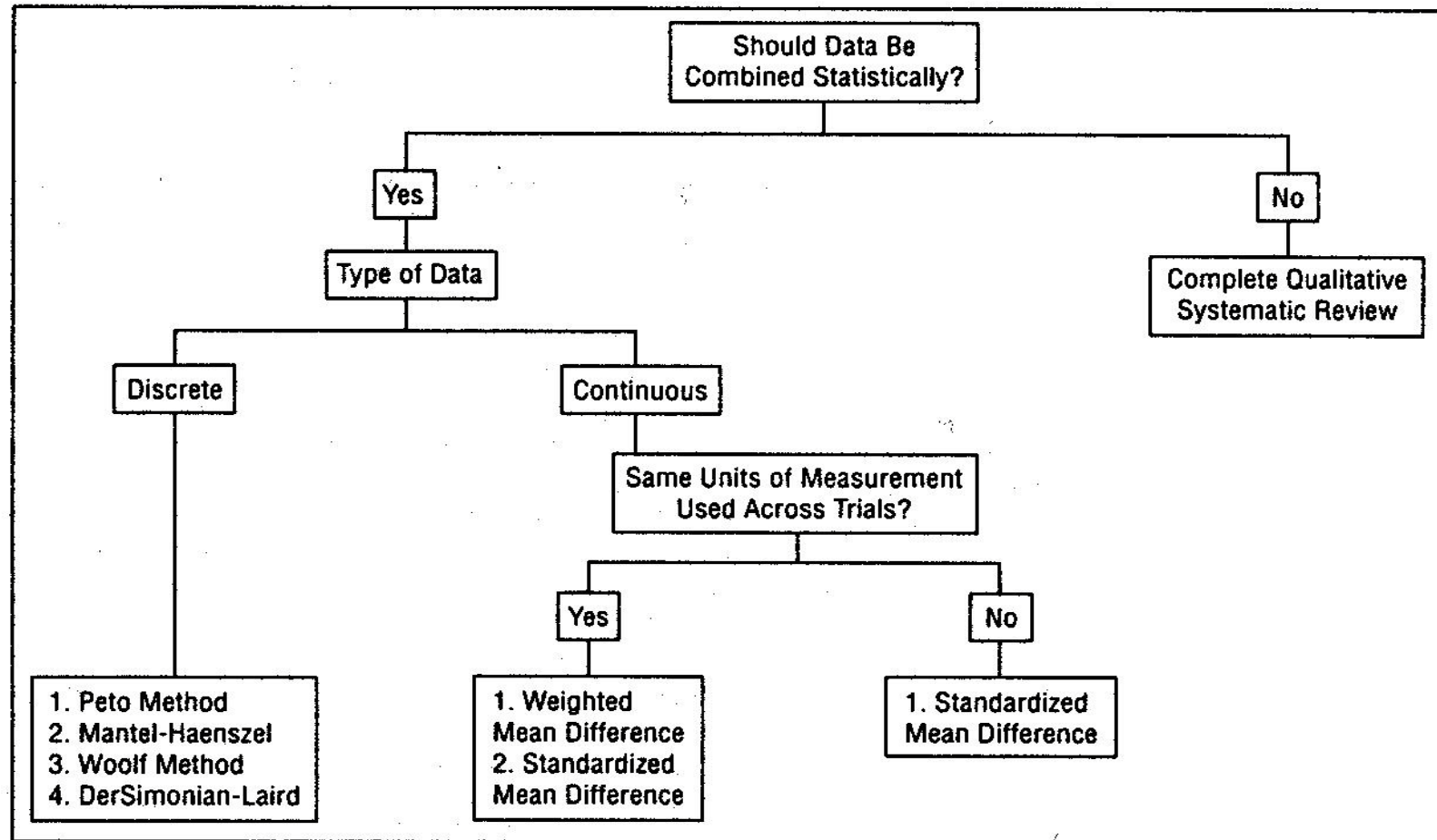
- ☛ Formulate the review question & write a protocol
- ☛ Search for and include primary studies
- ☛ Assess study quality
- ☛ Extract data
- ☛ Analyze data
- ☛ Interpret results & write a report

Road map for systematic reviews

Pai M, et al. *Natl Med J India* 2004;17(2):86-95.

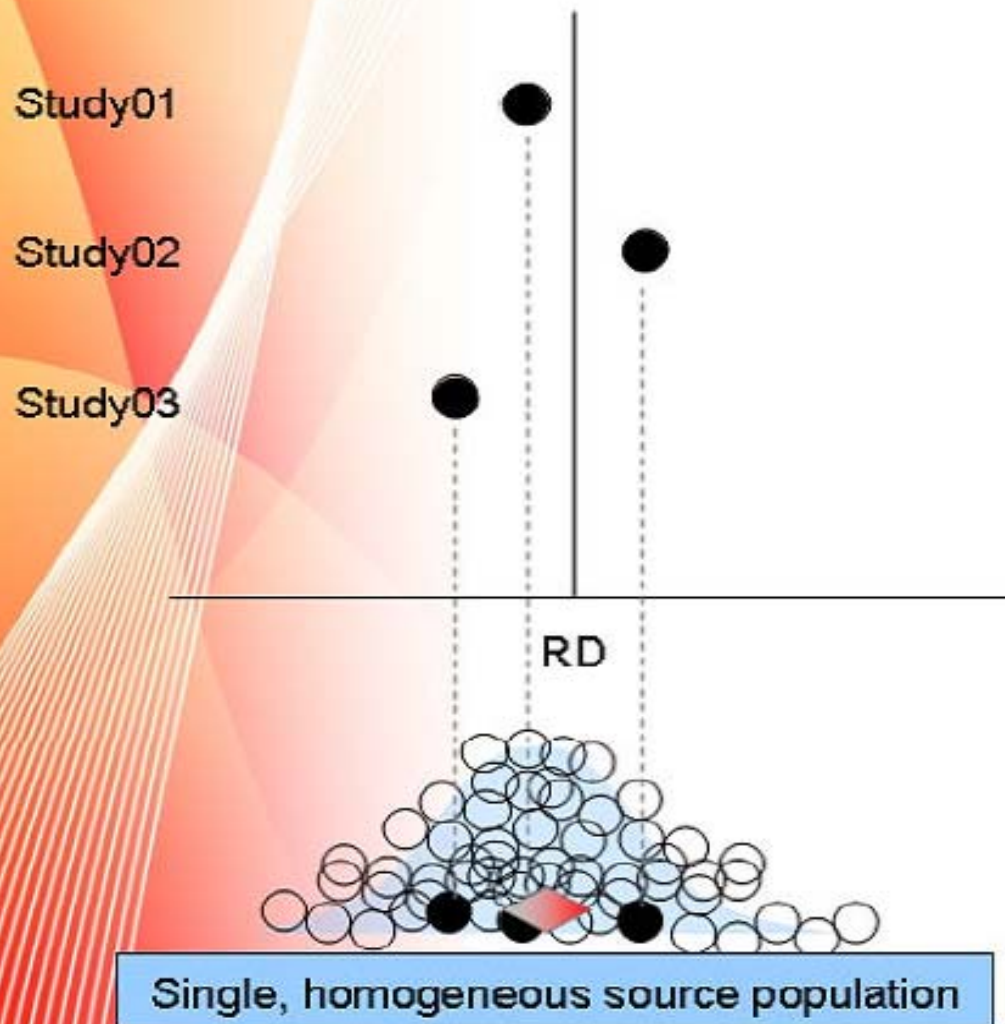


Meta-analysis options



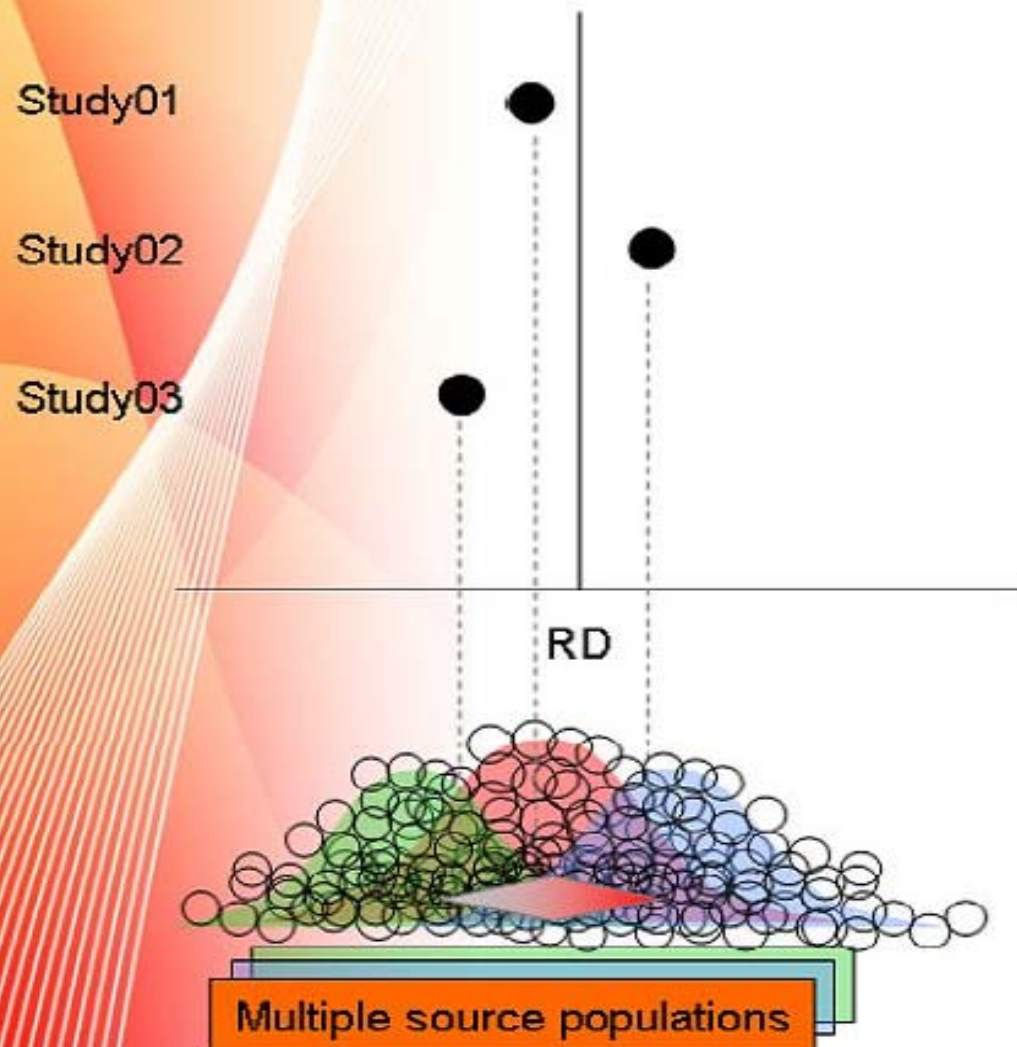
Algorithm of statistical choices available to systematic reviewers.

Fixed effect models



- In a fixed effect model, we assume the studies come from the same hypothetical population of studies
- We assume a single, 'fixed', parameter
- Study weights are a function of within-study variances
- Confidence interval relatively narrow

Random effects models



- There is no longer the assumption of a single, homogeneous source population
- Studies are allowed to come from different distributions
- Study weights not just based on within-study variances; a random effects constant is used to distribute the weights more evenly
- Results in a wider confidence interval

Forest Plot: when outcomes are dichotomous

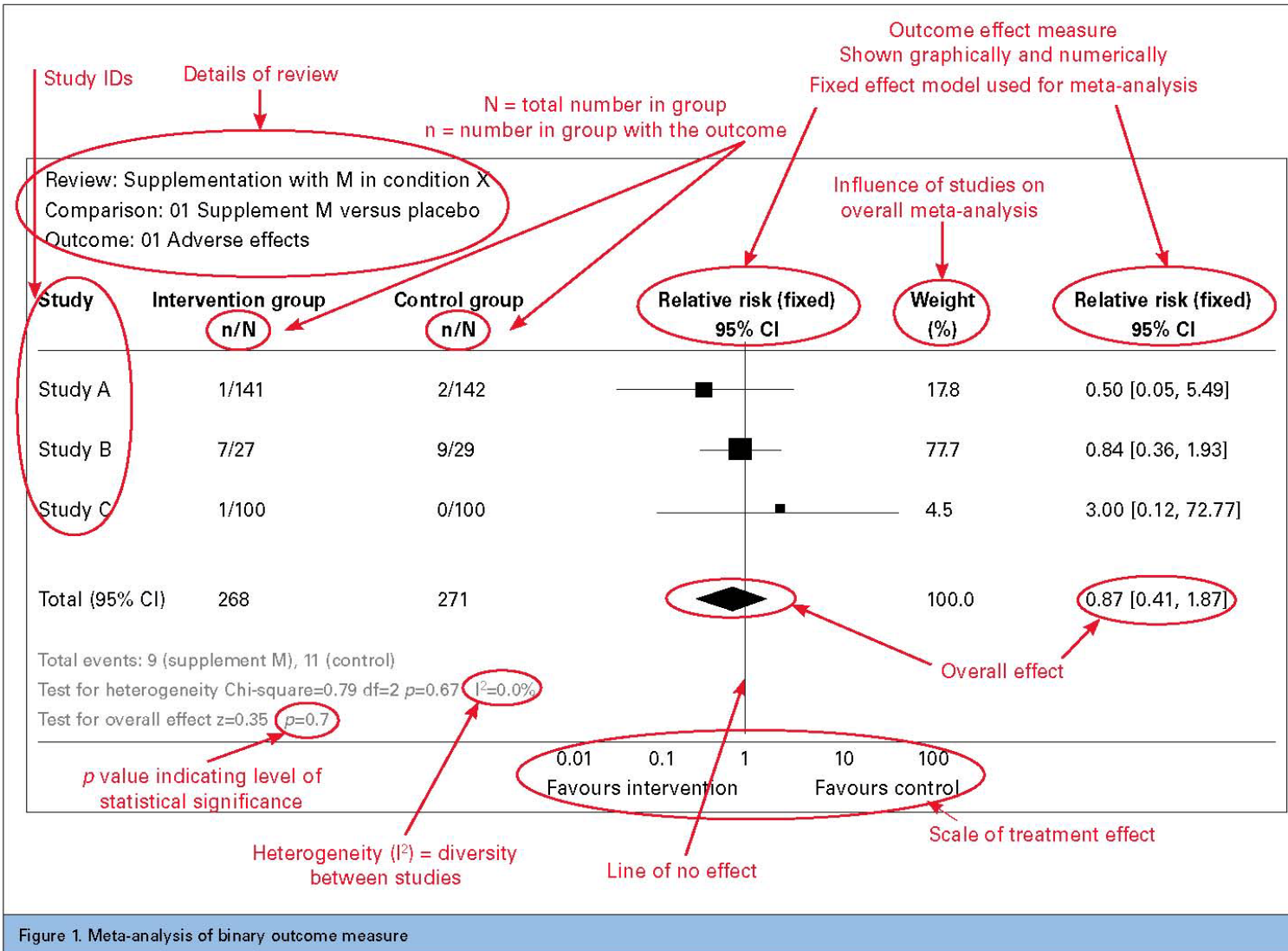
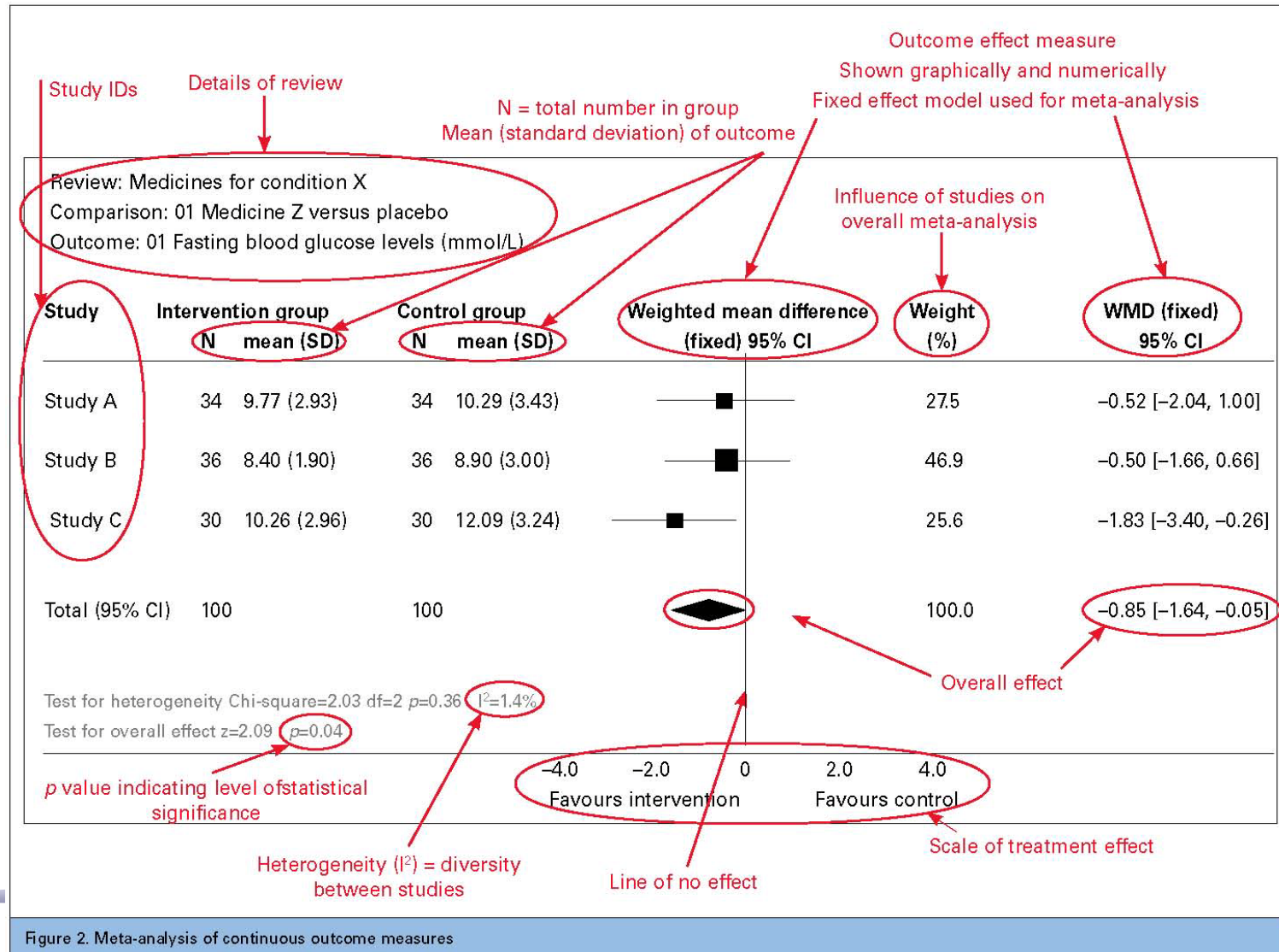
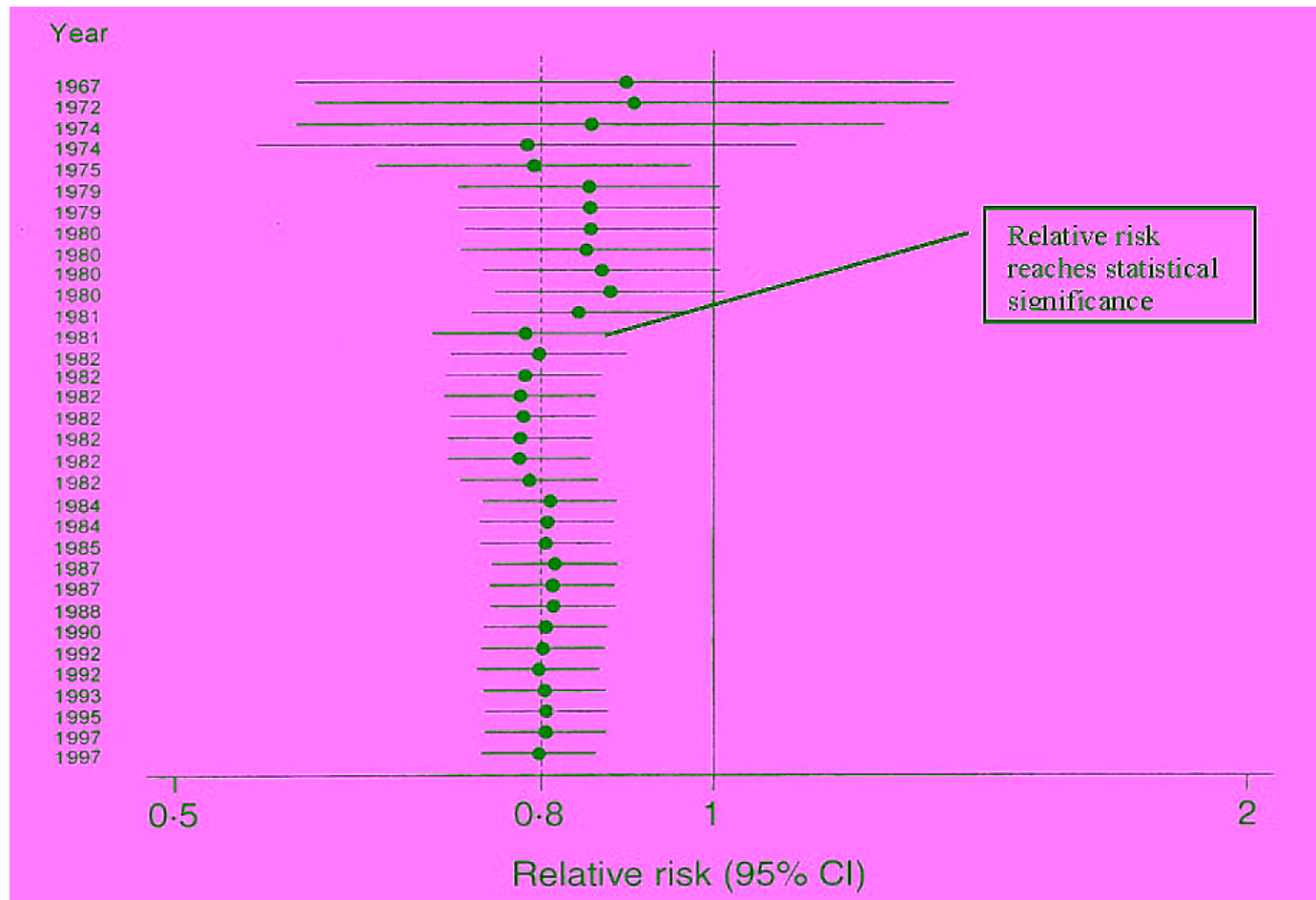


Figure 1. Meta-analysis of binary outcome measure

Forest Plot: when outcomes are continuous



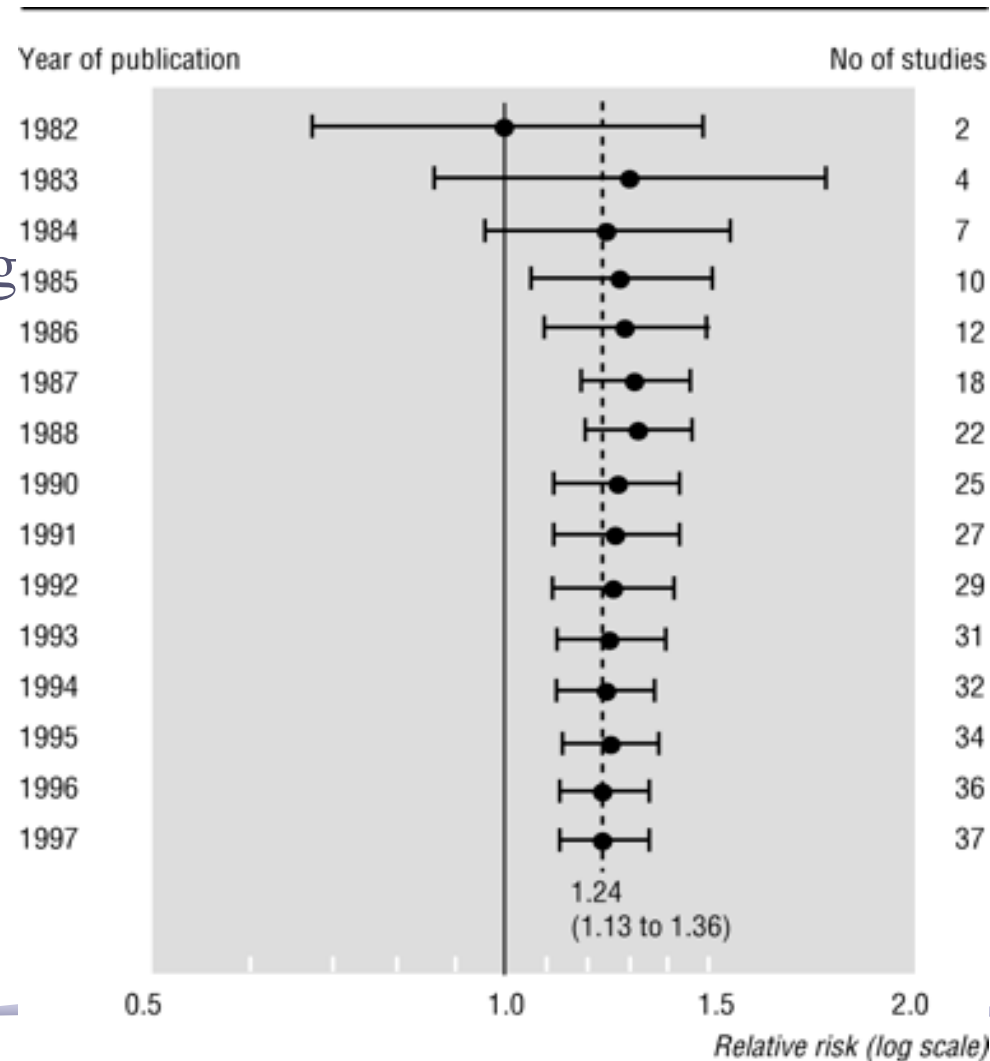
Forest Plot: Cumulative Meta-analysis



Beta-blockers after acute myocardial infarction

Cumulative Meta-analysis Plot

Passive smoking
and lung cancer
review



Meta-analysis Software



Free

- RevMan 5 [Review Manager]
- Meta-Analyst
- Epi Meta
- Easy MA
- Meta-DiSc
- Meta-Stat
- **MIX**

Commercial

- **Comprehensive Meta-analysis**
- Meta-Win
- WEasy MA

General stats packages (commercial)

- **Stata**
- SAS
- S-Plus

Meta-analysis demo

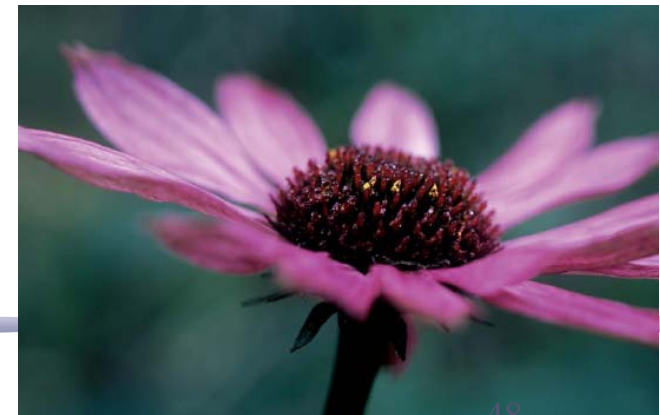
Evaluation of echinacea for the prevention and treatment of the common cold: a meta-analysis

Sachin A Shah, Stephen Sander, C Michael White, Mike Rinaldi, Craig I Coleman

Echinacea is one of the most commonly used herbal products, but controversy exists about its benefit in the prevention and treatment of the common cold. Thus, we did a meta-analysis evaluating the effect of echinacea on the incidence and duration of the common cold. 14 unique studies were included in the meta-analysis. Incidence of the common cold was reported as an odds ratio (OR) with 95% CI, and duration of the common cold was reported as the weighted mean difference (WMD) with 95% CI. Weighted averages and mean differences were calculated by a random-effects model (DerSimonian-Laird methodology). Heterogeneity was assessed by the Q statistic and review of L'Abbé plots, and publication bias was assessed through the Egger weighted regression statistic and visual inspection of funnel plots. Echinacea decreased the odds of developing the common cold by 58% (OR 0.42; 95% CI 0.25–0.71; Q statistic $p < 0.001$) and the duration of a cold by 1.4 days (WMD -1.44 , -2.24 to -0.64 ; $p = 0.01$). Similarly, significant reductions were maintained in subgroup analyses limited to Echinaguard/Echinacin use, concomitant supplement use, method of cold exposure, Jadad scores less than 3, or use of a fixed-effects model. Published evidence supports echinacea's benefit in decreasing the incidence and duration of the common cold.

Lancet Infect Dis 2007;7:473–80

University of Connecticut
School of Pharmacy, Storrs, CT,
USA (S A Shah PharmD,
S Sander PharmD,
C M White PharmD,
M Rinaldi PharmD,
C I Coleman PharmD);
University of Connecticut
Health Center Division of
Infectious Diseases,
Farmington, CT, USA
(M Rinaldi); and Division of
Drug Information, Hartford



	Analyses included in study	Incidence in echinacea group*	Incidence in control group*	Number of patients with cold in echinacea group	Number of patients with cold in control group	Mean duration in echinacea group (SD)	Mean duration in control group (SD)
Turner et al (2005) ¹⁴	Incidence of cold	73/149	58/103	NA	NA	NA	NA
Cohen et al (2004) ²⁸	Incidence of cold, duration of cold	85/160	150/168	138†	308†	1.60 (1.90)	2.90 (1.60)
Sperber et al (2004) ²²	Incidence of cold	14/24	18/22	NA	NA	NA	NA
Taylor et al (2003) ³²	Duration of cold	NA	NA	337†	370†	9.00 (9.37)	9.00 (9.81)
Barrett et al (2002) ²⁰	Duration of cold	NA	NA	69	73	6.27‡	5.75‡
Schulten et al (2001) ²³	Incidence of cold	35/41	38/39	NA	NA	NA	NA
Turner et al (2000) ³¹	Incidence of cold	11/50	14/42	NA	NA	NA	NA
Lindenmuth and Lindenmuth (2000) ²⁹	Duration of cold	NA	NA	48	47	2.34 (1.08)	4.33 (0.93)
Grimm and Muller (1999) ²⁴	Incidence of cold	35/54	40/54	NA	NA	NA	NA
Berg (1998) ²⁵	Incidence of cold	0/14	7/26	NA	NA	NA	NA
Melchart et al (1998) ¹⁵	Incidence of cold, duration of cold	60/199	33/90	60	33	8.00 (5.10)	8.7 (3.60)
Hoheisel et al (1997) ²⁷	Incidence of cold	24/60	36/60	NA	NA	NA	NA
Scaglione and Lund (1995) ³⁰	Duration of cold	NA	NA	16	16	3.37 (1.25)	4.37 (1.57)
Braunig and Knick (1993) ¹⁶	Duration of cold	NA	NA	70	45	9.10 (1.8)	12.9 (2.1)

NA=not applicable. *Data shown as number of events/total population. †Reported data is number of cold episodes, not number of patients with cold. ‡Reported data as difference of -0.52 days, 95% CI -1.09 to -0.22.

Table 2: Individual study characteristics

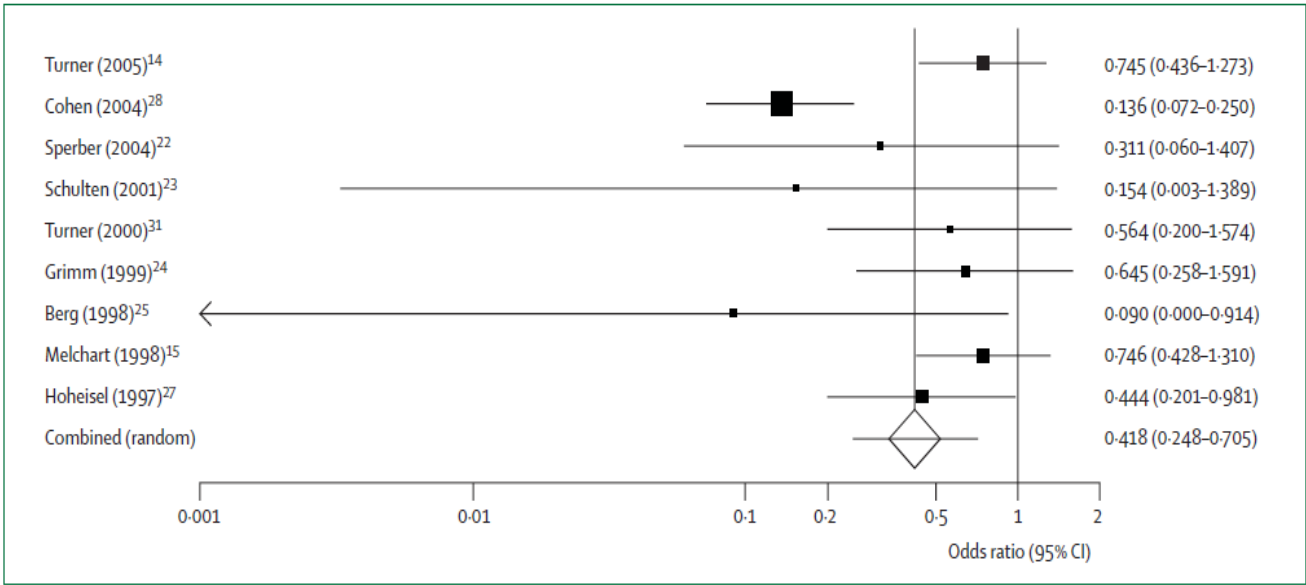
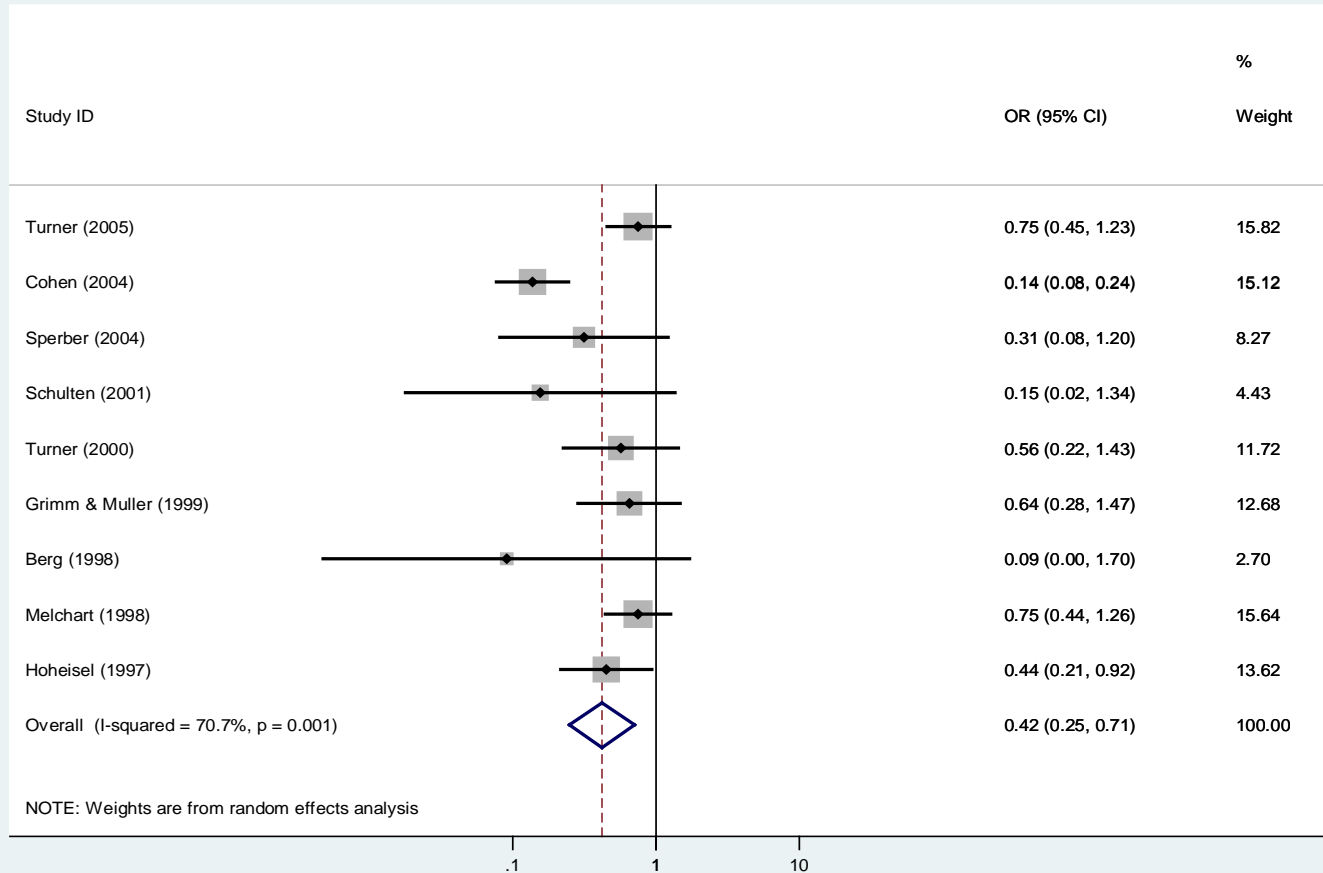


Figure 3: The effect of echinacea on incidence of common cold
The squares represent individual studies and the size of the square represents the weight given to each study in the meta-analysis. Error bars represent 95% CIs. The diamond represents the combined result. The solid vertical line extending upwards from 1.0 is the null value.

Meta-analysis using STATA



All systematic reviews are not meta-analyses!

- “...it is always appropriate and desirable to systematically review a body of data, but it may sometimes be inappropriate, or even misleading, to statistically pool results from separate studies. Indeed, it is our impression that reviewers often find it hard to resist the temptation of combining studies even when such meta-analysis is questionable or clearly inappropriate.”

- Egger et al. Systematic reviews in health care. London: BMJ books, 2001:5.

All systematic reviews are not systematic!

OPEN ACCESS Freely available online

PLOS MEDICINE

Editorial

Many Reviews Are Systematic but Some Are More Transparent and Completely Reported than Others

The PLoS Medicine Editors

OPEN ACCESS Freely available online

PLOS MEDICINE

Epidemiology and Reporting Characteristics of Systematic Reviews

David Moher^{1,2,3*}, Jennifer Tetzlaff¹, Andrea C. Tricco^{1,4}, Margaret Sampson¹, Douglas G. Altman⁵

1 Chalmers Research Group, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada, 2 Department of Paediatrics, Faculty of Medicine, University of Ottawa, Ottawa, Canada, 3 Department of Epidemiology and Community Medicine, Faculty of Medicine, University of Ottawa, Ottawa, Canada, 4 Institute of Population Health, University of Ottawa, Ottawa, Canada, 5 Centre for Statistics in Medicine, Oxford, United Kingdom

All systematic reviews are not systematic!

- 300 SRs were identified (one month)
- Majority (272 [90.7%]) reported in specialty journals
- Most reviews (213 [71.0%]) were categorized as therapeutic, and included a median of 16 studies
- Reviews typically searched a median of three electronic databases and two other sources
- Most (197/295 [66.8%]) reviews reported information about quality assessment, while few (68/294 [23.1%]) reported assessing for publication bias.
- A little over half (161/300 [53.7%]) reported combining their results statistically, of which most (147/161 [91.3%]) assessed for consistency across studies.
- There were large differences between Cochrane reviews and non-Cochrane reviews in the quality of reporting

When can meta-analyses mislead?

- ☛ When a meta-analysis is done outside of a systematic review
- ☛ When poor quality studies are included or when quality issues are ignored
- ☛ When small and inconclusive studies are included
- ☛ When inadequate attention is given to heterogeneity
 - Indiscriminate data aggregation can lead to inaccurate conclusions
- ☛ When reporting biases are a problem
 - Publication bias
 - Time lag bias
 - Duplicate publication bias
 - Language bias
 - Outcome reporting bias

SPECIAL ARTICLE

Selective Publication of Antidepressant Trials and Its Influence on Apparent Efficacy

Erick H. Turner, M.D., Annette M. Matthews, M.D., Eftihia Linardatos, B.S.,
Robert A. Tell, L.C.S.W., and Robert Rosenthal, Ph.D.

ABSTRACT

BACKGROUND

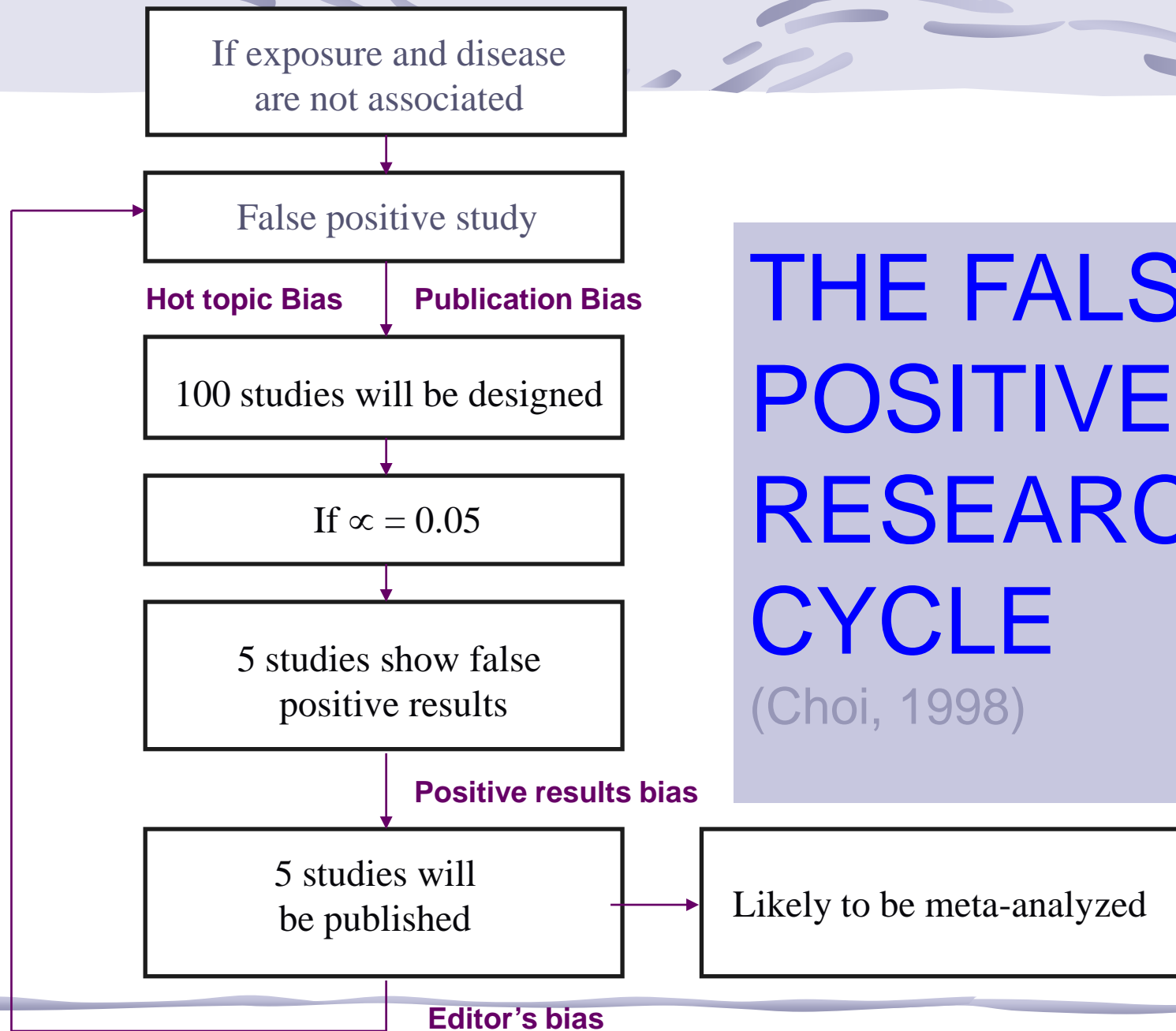
Evidence-based medicine is valuable to the extent that the evidence base is complete and unbiased. Selective publication of clinical trials — and the outcomes within those trials — can lead to unrealistic estimates of drug effectiveness and alter the apparent risk–benefit ratio.

METHODS

We obtained reviews from the Food and Drug Administration (FDA) for studies of 12 antidepressant agents involving 12,564 patients. We conducted a systematic literature search to identify matching publications. For trials that were reported in the literature, we compared the published outcomes with the FDA outcomes. We also compared the effect size derived from the published reports with the effect size derived from the entire FDA data set.

RESULTS

Among 74 FDA-registered studies, 31%, accounting for 3449 study participants, were not published. Whether and how the studies were published were associated with the study outcome. A total of 37 studies viewed by the FDA as having positive results were published; 1 study viewed as positive was not published. Studies viewed by the FDA as having negative or questionable results were, with 3 exceptions, either not published (22 studies) or published in a way that, in our opinion, conveyed a positive outcome (11 studies). According to the published literature, it appeared that 94% of the trials conducted were positive. By contrast, the FDA analysis showed that 51% were positive. Separate meta-analyses of the FDA and journal data sets showed that the increase in effect size ranged from 11 to 69% for individual drugs and was 32% overall.



THE FALSE POSITIVE RESEARCH CYCLE

(Choi, 1998)

Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials

Comparison of Protocols to Published Articles

An-Wen Chan, MD, DPhil
Asbjørn Hróbjartsson, MD, PhD
Mette T. Haahr, BSc
Peter C. Gøtzsche, MD, DrMedSci
Douglas G. Altman, DSc

SELECTIVE PUBLICATION OF STUDIES with statistically significant results has received widespread recognition.¹ In contrast, selective reporting of favorable outcomes within published studies has not undergone comparable empirical investigation. The existence of outcome reporting bias has been widely suspected for years,²⁻¹² but direct evidence is limited to case reports that have low generalizability¹³⁻¹⁵ and may themselves be subject to publication bias.

Our study had 3 goals: (1) to determine the prevalence of incomplete outcome reporting in published reports of randomized trials; (2) to assess the association between outcome reporting and statistical significance; and (3) to evaluate the consistency between primary outcomes specified in trial protocols and those defined in the published articles.

METHODS

Context Selective reporting of outcomes within published studies based on the nature or direction of their results has been widely suspected, but direct evidence of such bias is currently limited to case reports.

Objective To study empirically the extent and nature of outcome reporting bias in a cohort of randomized trials.

Design Cohort study using protocols and published reports of randomized trials approved by the Scientific-Ethical Committees for Copenhagen and Frederiksberg, Denmark, in 1994-1995. The number and characteristics of reported and unreported trial outcomes were recorded from protocols, journal articles, and a survey of trialists. An outcome was considered incompletely reported if insufficient data were presented in the published articles for meta-analysis. Odds ratios relating the completeness of outcome reporting to statistical significance were calculated for each trial and then pooled to provide an overall estimate of bias. Protocols and published articles were also compared to identify discrepancies in primary outcomes.

Main Outcome Measures Completeness of reporting of efficacy and harm outcomes and of statistically significant vs nonsignificant outcomes; consistency between primary outcomes defined in the most recent protocols and those defined in published articles.

Results One hundred two trials with 122 published journal articles and 3736 outcomes were identified. Overall, 50% of efficacy and 65% of harm outcomes per trial were incompletely reported. Statistically significant outcomes had a higher odds of being fully reported compared with nonsignificant outcomes for both efficacy (pooled odds ratio, 2.4; 95% confidence interval [CI], 1.4-4.0) and harm (pooled odds ratio, 4.9; 95% CI, 1.8-12.0) data. In comparing published articles with protocols, 62% of trials had at least 1 primary outcome that was changed, introduced, or omitted. Eighty percent of survey responders (42/49) denied the existence of unreported outcomes despite clear evidence to the contrary.

Conclusions The reporting of trial outcomes is not only frequently incomplete but also biased and inconsistent with protocols. Published articles, as well as reviews that incorporate them, may therefore be unreliable and overestimate the benefits of intervention. To ensure transparency, planned trials should be registered and protocols should be made publicly available prior to trial completion.

Selection in Reported Epidemiological Risks: An Empirical Assessment

Fotini K. Kavvoura¹, George Liberopoulos¹, John P. A. Ioannidis^{1,2*}

¹ Clinical and Molecular Epidemiology Unit, Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece, ² Department of Medicine, Tufts University School of Medicine, Boston, Massachusetts, United States of America

Funding: The authors received no specific funding for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Academic Editor: Eduardo L. Franco, McGill University, Canada

Citation: Kavvoura FK, Liberopoulos G, Ioannidis JPA (2007) Selection in reported epidemiological risks: An empirical assessment. *PLoS Med* 4(3): e79. doi:10.1371/journal.pmed.0040079

Received: August 22, 2006
Accepted: January 8, 2007
Published: March 6, 2007

Copyright: © 2007 Kavvoura et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: ANOVA, analysis of variance; CI, confidence interval; IQR, interquartile range

* To whom correspondence should be addressed. E-mail: joannidis@cc.uoi.gr

ABSTRACT

Background

Epidemiological studies may be subject to selective reporting, but empirical evidence thereof is limited. We empirically evaluated the extent of selection of significant results and large effect sizes in a large sample of recent articles.

Methods and Findings

We evaluated 389 articles of epidemiological studies that reported, in their respective abstracts, at least one relative risk for a continuous risk factor in contrasts based on median, tertile, quartile, or quintile categorizations. We examined the proportion and correlates of reporting statistically significant and nonsignificant results in the abstract and whether the magnitude of the relative risks presented (coined to be consistently ≥ 1.00) differs depending on the type of contrast used for the risk factor. In 342 articles (87.9%), ≥ 1 statistically significant relative risk was reported in the abstract, while only 169 articles (43.4%) reported ≥ 1 statistically nonsignificant relative risk in the abstract. Reporting of statistically significant results was more common with structured abstracts, and was less common in US-based studies and in cancer outcomes. Among 50 randomly selected articles in which the full text was examined, a median of nine (interquartile range 5–16) statistically significant and six (interquartile range 3–16) statistically nonsignificant relative risks were presented ($p = 0.25$). Paradoxically, the smallest presented relative risks were based on the contrasts of extreme quintiles; on average, the relative risk magnitude was 1.41-, 1.42-, and 1.36-fold larger in contrasts of extreme quartiles, extreme tertiles, and above-versus-below median values, respectively ($p < 0.001$).

Conclusions

Published epidemiological investigations almost universally highlight significant associations between risk factors and outcomes. For continuous risk factors, investigators selectively present contrasts between more extreme groups, when relative risks are inherently lower.

JAMA 2004

PLoS Med 2007

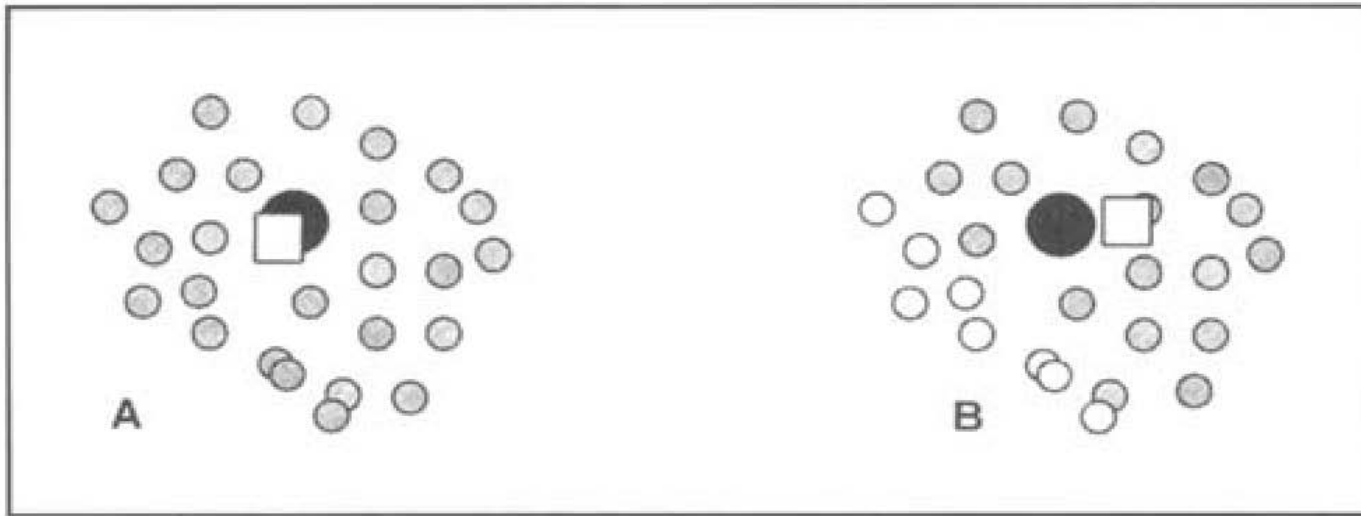
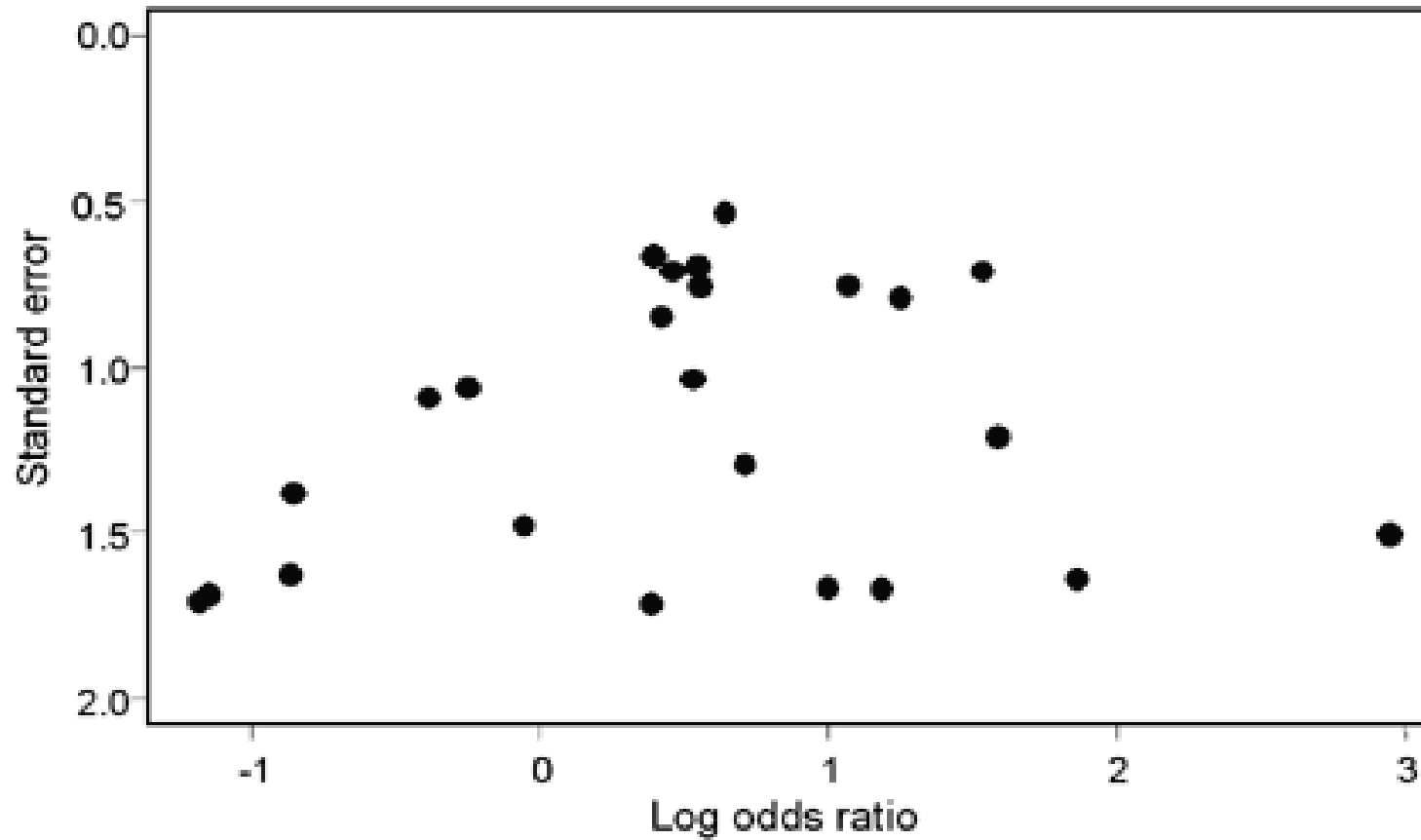
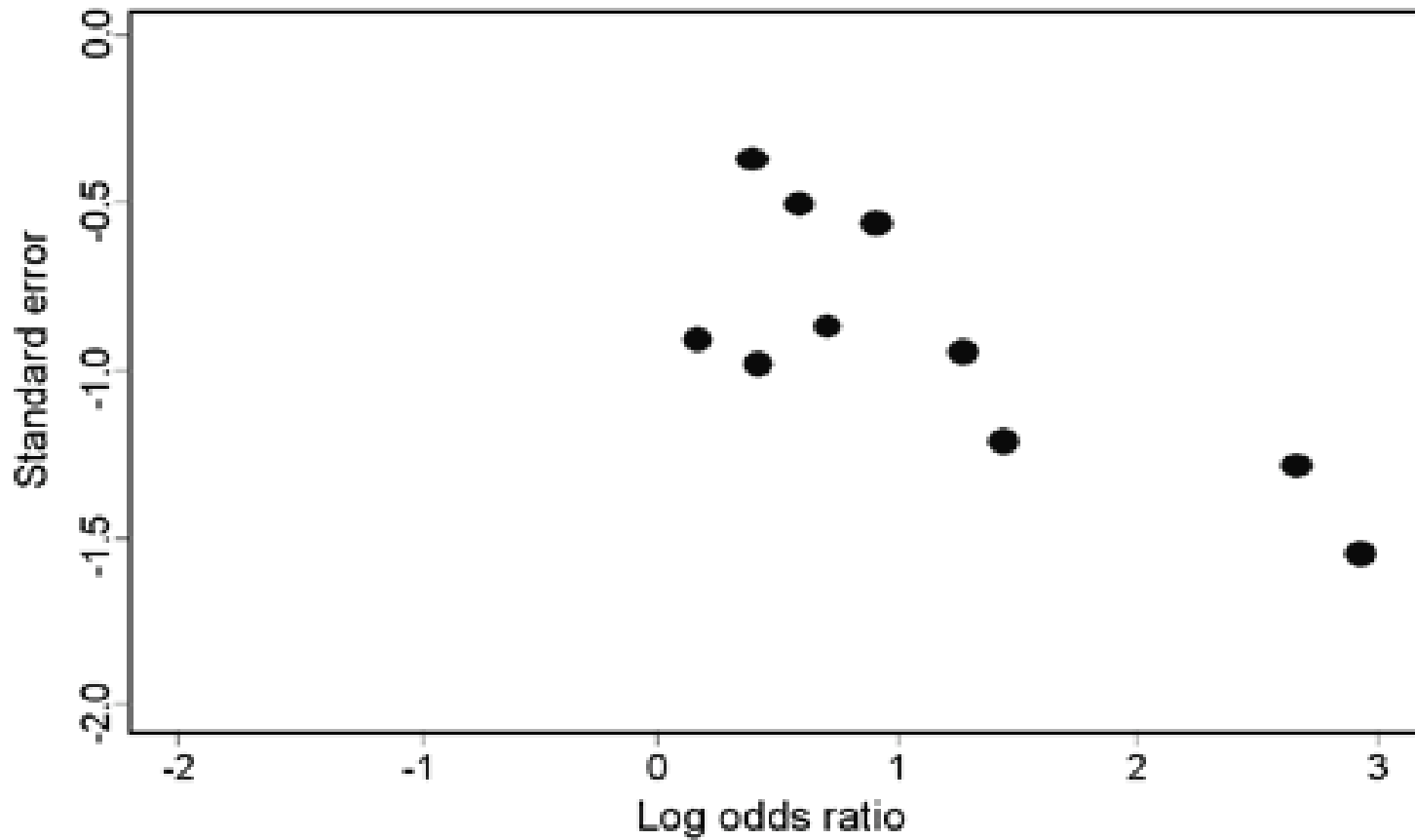


Figure 1. Publication bias. A, The black circle represents the underlying truth. The white square represents the pooled estimate from a systematic review of all the evidence (small shaded circles). B, The white circles represent evidence that was not identified by the reviewers because it was not published. Note the error in the pooled estimate (publication bias).

Funnel plot to detect publication bias



Funnel plot to detect publication bias



Contradicted and Initially Stronger Effects in Highly Cited Clinical Research

John P. A. Ioannidis, MD

CLINICAL RESEARCH ON IMPORTANT questions about the efficacy of medical interventions is sometimes followed by subsequent studies that either reach opposite conclusions or suggest that the original claims were too strong. Such disagreements may upset clinical practice and acquire publicity in both scientific circles and in the lay press. Several empirical investigations have tried to address whether specific types of studies are more likely to be contradicted and to explain observed controversies. For example, evidence exists that small studies may sometimes be refuted by larger ones.^{1,2}

Similarly, there is some evidence on disagreements between epidemiological studies and randomized trials.³⁻⁵ Prior investigations have focused on a variety of studies without any particular attention to their relative importance and scientific impact. Yet, most research publications have little impact while a small minority receives

Context Controversy and uncertainty ensue when the results of clinical research on the effectiveness of interventions are subsequently contradicted. Controversies are most prominent when high-impact research is involved.

Objectives To understand how frequently highly cited studies are contradicted or find effects that are stronger than in other similar studies and to discern whether specific characteristics are associated with such refutation over time.

Design All original clinical research studies published in 3 major general clinical journals or high-impact-factor specialty journals in 1990-2003 and cited more than 1000 times in the literature were examined.

Main Outcome Measure The results of highly cited articles were compared against subsequent studies of comparable or larger sample size and similar or better controlled designs. The same analysis was also performed comparatively for matched studies that were not so highly cited.

Results Of 49 highly cited original clinical research studies, 45 claimed that the intervention was effective. Of these, 7 (16%) were contradicted by subsequent studies, 7 others (16%) had found effects that were stronger than those of subsequent studies, 20 (44%) were replicated, and 11 (24%) remained largely unchallenged. Five of 6 highly-cited nonrandomized studies had been contradicted or had found stronger effects vs 9 of 39 randomized controlled trials ($P = .008$). Among randomized trials, studies with contradicted or stronger effects were smaller ($P = .009$) than replicated or unchallenged studies although there was no statistically significant difference in their early or overall citation impact. Matched control studies did not have a significantly different share of refuted results than highly cited studies, but they included more studies with "negative" results.

Conclusions Contradiction and initially stronger effects are not unusual in highly cited research of clinical interventions and their outcomes. The extent to which high citations may provoke contradictions and vice versa needs more study. Controversies are most common with highly cited nonrandomized studies, but even the most highly cited randomized trials may be challenged and refuted over time, especially small ones.

JAMA. 2005;294:218-228

www.jama.com

Discrepancies between meta-analyses

A guide to interpreting discordant systematic reviews

Alejandro R. Jadad, MD, DPhil; Deborah J. Cook, MD, MSc;
George P. Browman, MD, MSc

Abstract

SYSTEMATIC REVIEWS ARE BECOMING prominent tools to guide health care decisions. As the number of published systematic reviews increases, it is common to find more than 1 systematic review addressing the same or a very similar therapeutic question. Despite the promise for systematic reviews to resolve conflicting results of primary studies, conflicts among reviews are now emerging. Such conflicts produce difficulties for decision-makers (including clinicians, policy-makers, researchers and patients) who rely on these reviews to help them make choices among alternative interventions when experts and the results of trials disagree. The authors provide an adjunct decision tool — a decision algorithm — to help decision-makers select from among discordant reviews.

Résumé

LES EXAMENS CRITIQUES SYSTÉMATIQUES DEVIENNENT des outils importants pour guider les décisions relatives aux soins de santé. Comme le nombre des examens critiques



Education

Éducation

Dr. Jadad is with the Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, and Drs. Cook and Browman are with the Departments of Clinical Epidemiology and Biostatistics and of Medicine, McMaster University, Hamilton, Ont.

This article has been peer reviewed.

Can Med Assoc J 1997;156:1411-6

Discrepancies between meta-analyses and mega-trials

The New England Journal of Medicine

Special Article

DISCREPANCIES BETWEEN META-ANALYSES AND SUBSEQUENT LARGE RANDOMIZED, CONTROLLED TRIALS

JACQUES LELORIER, M.D., PH.D., GENEVIÈVE GRÉGOIRE, M.D., ABDELTIF BENHADDAD, M.D., JULIE LAPIERRE, M.D.,
AND FRANÇOIS DERDERIAN, M.Sc.



Available online at www.sciencedirect.com



Contemporary Clinical Trials 28 (2007) 324–328

Contemporary
Clinical
Trials

www.elsevier.com/locate/conclintrial

Discussion

Mega-trials vs. meta-analysis: Precision vs. heterogeneity?

Ian Shrier^{a,*}, Robert W. Platt^b, Russell J. Steele^c

Yes, there are problems, but

- meta-analysis has made and continues to make major contributions to medical research, clinical decision making, and standards of research reportage. However, it is no panacea. Readers need to examine any meta-analyses critically to see whether researchers have overlooked important sources of clinical heterogeneity among the included trials. They should demand evidence that the authors undertook a comprehensive search, avoiding covert duplicate data and unearthing unpublished trials and data. Lastly, readers and researchers alike need to appreciate that not every systematic review should lead to an actual meta-analysis...

David Naylor. BMJ 1997;315:617-619



McGill Summer Session in Epidemiology and Biostatistics 2010

The Summer Session in Epidemiology and Biostatistics at McGill offers health professionals the opportunity to gain familiarity with the principles of epidemiology and biostatistics. It also offers graduate students from McGill and other universities the opportunity to acquire academic credits and thereby accelerate course work during a summer term. Summer session website: <http://www.mcgill.ca/epi-biostat-occh/summer/>

EPIB-672: Systematic Reviews and Meta-analyses

Special topics in Epidemiology and Biostatistics

Academic credits: 2

Dates: May 3 to 14, 2010

Class times: 2.00 – 4.30 PM, Monday through Friday

Faculty: Dr. Madhukar Pai, MD, PhD (madhukar.pai@mcgill.ca)

Enrollment limit: 25

Systematic reviews and meta-analyses are critical for evidence-based clinical and public health practice. The widespread and growing application of systematic reviews to synthesize evidence on key research and clinical questions makes it useful for health professionals to be able to understand and critique this research design. This course will provide a detailed description of the systematic review process, discuss the strengths and limitations of the method, and provide step-by-step guidance on how to actually perform a systematic review.

Types of review articles



Specific topics to be covered include: formulation of the review question, searching of literature, quality assessment of studies, data extraction, meta-analytic methods, and report writing. The course will also cover statistical issues such as selection of statistical models for meta-analysis, practical examples of fixed and random effects models as well as examples of methods to evaluate heterogeneity and publication bias; graphical and tabular templates for the presentation of meta-analysis data. STATA software package will be used, along with computer lab tutorials on how to effectively use tools such as PubMed and EndNote for conducting reviews. This course will feature invited speakers who will provide overviews of special topics.

Prerequisites: Introductory level training in epidemiology (e.g. EPIB601) and biostatistics (e.g. EPIB607), or permission of the instructor. Students who have not done prior coursework in introductory epidemiology/biostatistics must contact the instructor.

Recommended textbook: Egger M, Smith GD, Altman DG (editors). Systematic reviews in health care. Meta-analysis in context. Second Edition. London: BMJ Publishing Group, 2001. URL: www.systematicreviews.com. All participants will receive a CD that contains useful resources for systematic reviews, such as free software, guidelines, sample data extraction forms, quality checklists, powerpoint slides, etc.

Note: The language of instruction is English, and students are advised that fluency in English is essential to benefit from the course. However, students may submit their course assignments and examinations in French. Courses may be taken for Academic Credit, Continuing Medical Education (CME) Credit, or for a Professional Interest Certificate.

