Assessment of Study Quality and Data Extraction

Madhukar Pai, MD, PhD
Associate Professor
Department of Epidemiology & Biostatistics
McGill University, Montreal, Canada
Email: madhukar.pai@mcgill.ca

Quality assessment

- The rationale
  - The quality of evidence generated by a review depends entirely on the quality of primary studies which make up the review
  - A meta-analysis of poor quality studies can not generate high quality summary results!
    - "…even elegant statistical manipulations, when performed on biased rubble, are incapable of generating unbiased precious stones." – Iain Chalmers
  - Quality assessment is a vital component of all systematic reviews
    - sets apart a systematic review from a narrative review
  - Many reviews which claim to be 'systematic reviews' do not perform quality assessment

Quality assessment

- What does quality mean?
  - Quality means different things to different people
  - In the context of systematic reviews, quality refers to methodological quality – the internal validity of primary studies
  - Internal validity = lack of bias
  - Bias
    - Selection bias
    - Information bias
    - Confounding

Guidelines and recommendations: GRADE

Box 2 | Quality of evidence and definitions

High quality— Further research is very unlikely to change our confidence in the estimate of effect

Moderate quality— Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate

Low quality— Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate

Very low quality— Any estimate of effect is very uncertain

Factors in deciding on quality of evidence

Factors that might decrease quality of evidence

- Study limitations
- Inconsistency of results
- Indirectness of evidence
- Imprecision
- Publication bias
- Factors that might increase quality of evidence
- Large magnitude of effect
- Plausible confounding, which would reduce a demonstrated effect
- Dose-response gradient

http://www.gradeworkinggroup.org/

Quality assessment

- Why assess quality?
  - To set minimum quality standards for inclusion of studies
    - Approaches:
      - Include only high quality studies in the review (e.g. Cochrane mammography review)
      - Include all studies and ignore quality
      - Include all studies and then look at the impact of quality on the study results (using subgroup analysis, meta-regression, etc.)

Case study: "Same trials, different takes"

- Mammography for breast cancer is an established screening method
- Is screening with mammography justifiable?
- Gotzsche & Olsen [Nordic Cochrane Centre] conducted a systematic review in 2000 and updated it in 2001.
- They identified 8 large RCTs on this topic, with over 182,000 women randomized

Gotzsche & Olsen. Is screening for breast cancer with mammography justifiable? Lancet 2000;355:129 – 34.
Olsen & Gotzsche. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340-42.

"Same trials, different takes"

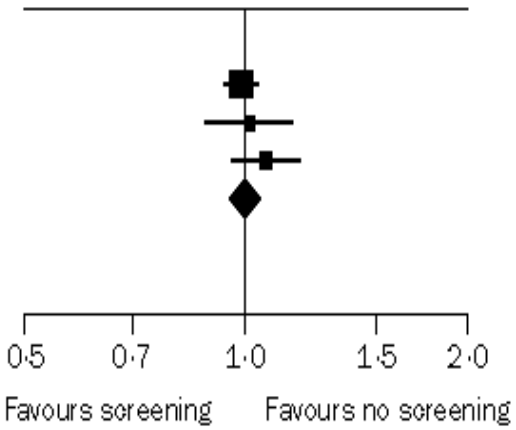
- The authors found that no trial data were of high quality
  - Two were of medium quality, and the rest were poor quality or flawed.
- When the results of the two medium quality trials were combined, the risk ratio was 1.00 (95% CI 0.96, 1.05)
- They concluded that "screening for breast cancer with mammography is unjustified" and "any hope or claim that screening mammography with more modern technologies than applied in these trials will reduce mortality without causing too much harm will have to be tested in large, well-conducted randomised trials…"

Gotzsche & Olsen. Is screening for breast cancer with mammography justifiable? Lancet 2000;355:129 – 34.
Olsen & Gotzsche. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340-42.

"Same trials, different takes"

Study	Screened Number of deaths/ number of women	Not screened Number of deaths/ number of women	Relative risk* (95% CI)	Weight (%)	Relative risk* (95% CI)
Malmö 1976	2537/21088	2593/21195		70·08	0·98 (0·93–1·04)
Canada 1980a	418/25214	414/25216		11·22	1·01 (0·88–1·16)
Canada 1980b	734/19711	690/19694		18·70	1·06 (0·96–1·18)
Subtotal	3689/66013	3697/66105		100·00	1·00 (0·96–1·05)

Test for heterogeneity: χ^2=1·80, df=2 (p=0·41)
Test for overall effect: z=0·05 (p=0·96)



All-cause mortality in medium-quality screening trials after 13 years

*Fixed-effects model.

Olsen & Gotzsche. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340-42.

"Same trials, different takes"

- The US Preventive Services Task Force (2002) used the same set of 8 trials:
  - "Recently, a 2001 Cochrane Collaboration review of the same trials concluded that six of the eight trials were "flawed" or of "poor quality" and that the pooled results from the remaining two better trials did not support a benefit from mammography.
  - Although the USPSTF was concerned about many (but not all) of the flaws identified in this review, it did not consider the presence of flaws sufficient reason in itself for rejecting trial results.
  - The meta-analysis performed for the USPSTF on the most current published data found that the pooled effect size of the combined trials was sizable and statistically significant: the summary relative risk (RR) of breast cancer death among women randomized to screening in seven trials that included women older than 50 was 0.77 (95 percent CI, 0.67-0.89).
  - *The USPSTF recommends screening mammography, with or without clinical breast examination, every 1-2 years for women aged 40 and older."*

http://www.ahcpr.gov/clinic/3rduspstf/breast cancer/

2009 Update of the USPSTF guidelines

CLINICAL GUIDELINES | Annals of Internal Medicine

Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement

U.S. Preventive Services Task Force*

Description: Update of the 2002 U.S. Preventive Services Task Force (USPSTF) recommendation statement on screening for breast cancer in the general population.

Methods: The USPSTF examined the evidence on the efficacy of 5 screening modalities in reducing mortality from breast cancer: film mammography, clinical breast examination, breast self-examination, digital mammography, and magnetic resonance imaging in order to update the 2002 recommendation. To accomplish this update, the USPSTF commissioned 2 studies: 1) a targeted systematic evidence review of 6 selected questions relating to benefits and harms of screening, and 2) a decision analysis that used population modeling techniques to compare the expected health outcomes and resource requirements of starting and ending mammography screening at different ages and using annual versus biennial screening intervals.

Recommendations: The USPSTF recommends against routine screening mammography in women aged 40 to 49 years. The decision to start regular, biennial screening mammography before the age of 50 years should be an individual one and take into account patient context, including the patient's values regarding specific benefits and harms. (Grade C recommendation)

The USPSTF recommends biennial screening mammography for women between the ages of 50 and 74 years. (Grade B recommendation)

The USPSTF concludes that the current evidence is insufficient to assess the additional benefits and harms of screening mammography in women 75 years or older. (I statement)

The USPSTF concludes that the current evidence is insufficient to assess the additional benefits and harms of clinical breast examination beyond screening mammography in women 40 years or older. (I statement)

The USPSTF recommends against clinicians teaching women how to perform breast self-examination. (Grade D recommendation)

The USPSTF concludes that the current evidence is insufficient to assess additional benefits and harms of either digital mammography or magnetic resonance imaging instead of film mammography as screening modalities for breast cancer. (I statement)

Ann Intern Med. 2009;151:716-726. www.annals.org

For author affiliation, see end of text.

* For a list of the members of the USPSTF, see the **Appendix** (available at www.annals.org).

Quality assessment

- Why assess quality?
  - To see if differences in study quality can explain heterogeneity in study results
  - To weight the study results by quality
    - Higher quality studies are weighted more than lower quality studies
  - To allow the readers to interpret the strength of evidence
  - To identify future research questions

Quality assessment

- Quality assessment instruments
  - Score approach using quality scales
    - Each item in a scale gets a numeric score
    - An overall quality score is generated by adding up the scores of each item
    - Cut-offs are then used to categorize as “high” vs “low” quality
  - Component approach
    - Items in a checklist are scored as Yes or No
    - No overall numeric score is computed
    - Subjective judgment on high vs low quality

The Hazards of Scoring the Quality of Clinical Trials for Meta-analysis

Peter Jüni, MD
Anne Witschi, MD
Ralph Bloch, MD, PhD
Matthias Egger, MD, MSc

Context Although it is widely recommended that clinical trials undergo some type of quality review, the number and variety of quality assessment scales that exist make it unclear how to achieve the best assessment.

Objective To determine whether the type of quality assessment scale used affects the conclusions of meta-analytic studies.

Design and Setting Meta-analysis of 17 trials comparing low-molecular-weight heparin (LMWH) with standard heparin for prevention of postoperative thrombosis using 25 different scales to identify high-quality trials. The association between treatment effect and summary scores and the association with 3 key domains (concealment of treatment allocation, blinding of outcome assessment, and handling of withdrawals) were examined in regression models.

Main Outcome Measure Pooled relative risks of deep vein thrombosis with LMWH vs standard heparin in high-quality vs low-quality trials as determined by 25 quality scales.

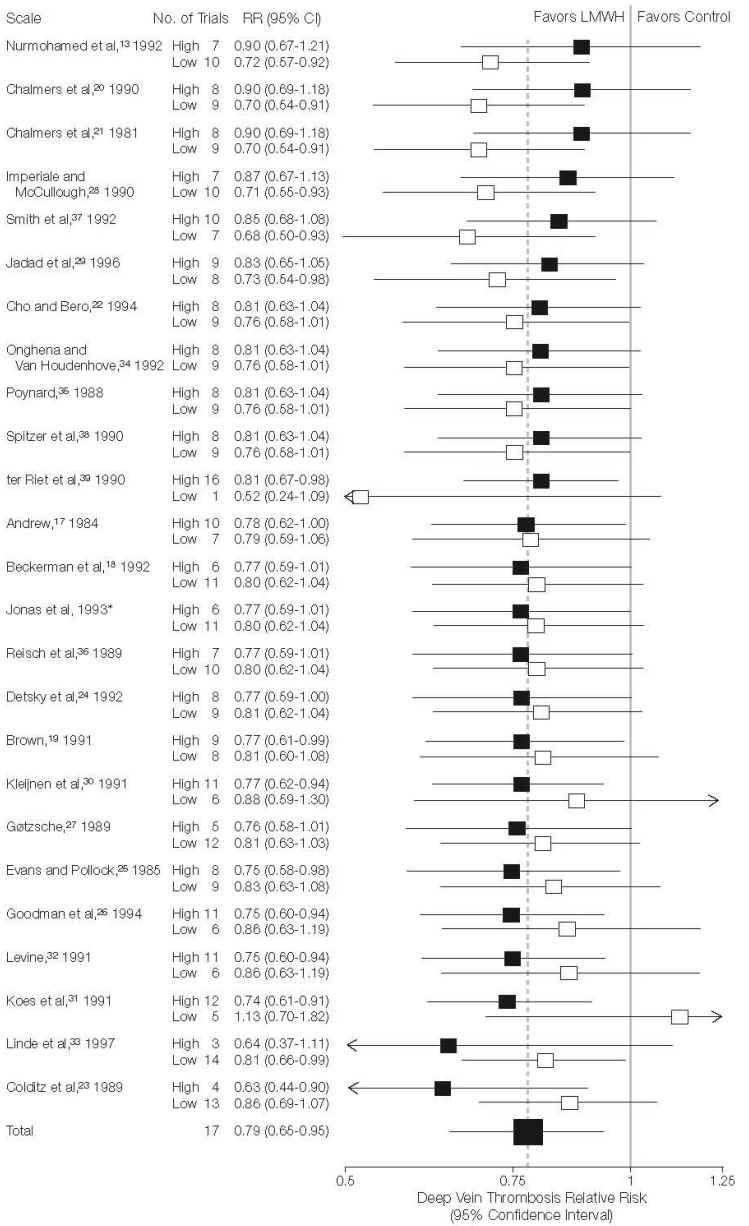
Results Pooled relative risks from high-quality trials ranged from 0.63 (95% confidence interval [CI], 0.44-0.90) to 0.90 (95% CI, 0.67-1.21) vs 0.52 (95% CI, 0.24-1.09) to 1.13 (95% CI, 0.70-1.82) for low-quality trials. For 6 scales, relative risks of high-quality trials were close to unity, indicating that LMWH was not significantly superior to standard heparin, whereas low-quality trials showed better protection with LMWH ($P<.05$). Seven scales showed the opposite: high quality trials showed an effect whereas low quality trials did not. For the remaining 12 scales, effect estimates were similar in the 2 quality strata. In regression analysis, summary quality scores were not significantly associated with treatment effects. There was no significant association of treatment effects with allocation concealment and handling of withdrawals. Open outcome assessment, however, influenced effect size with the effect of LMWH, on average, being exaggerated by 35% (95% CI, 1%-57%; $P=.046$).

Conclusions Our data indicate that the use of summary scores to identify trials of high quality is problematic. Relevant methodological aspects should be assessed individually and their influence on effect sizes explored.

JAMA. 1999;282:1054-1060 www.jama.com

ALTHOUGH RANDOMIZED CONtrolled trials provide the best evidence of the efficacy of medical interventions, they are not immune to bias. Studies relating methodological features of trials to their results have shown that trial quality influences effect sizes. For populations of trials examining treatments in myocardial infarction,[1] perinatal medicine,[2] and various disease areas,[3] it has consistently been shown that inadequate concealment of treatment allocation, resulting, for example, from the use of open random-number tables, is associated on average with larger treatment effects. One of these studies[2] also found larger average effect sizes if trials were not double-blind. Analyses of individual trials suggest that in some instances effect sizes are also overestimated if some participants, for example, those not adhering to study medications, were excluded from the analysis.[4-6] Informal qualita-

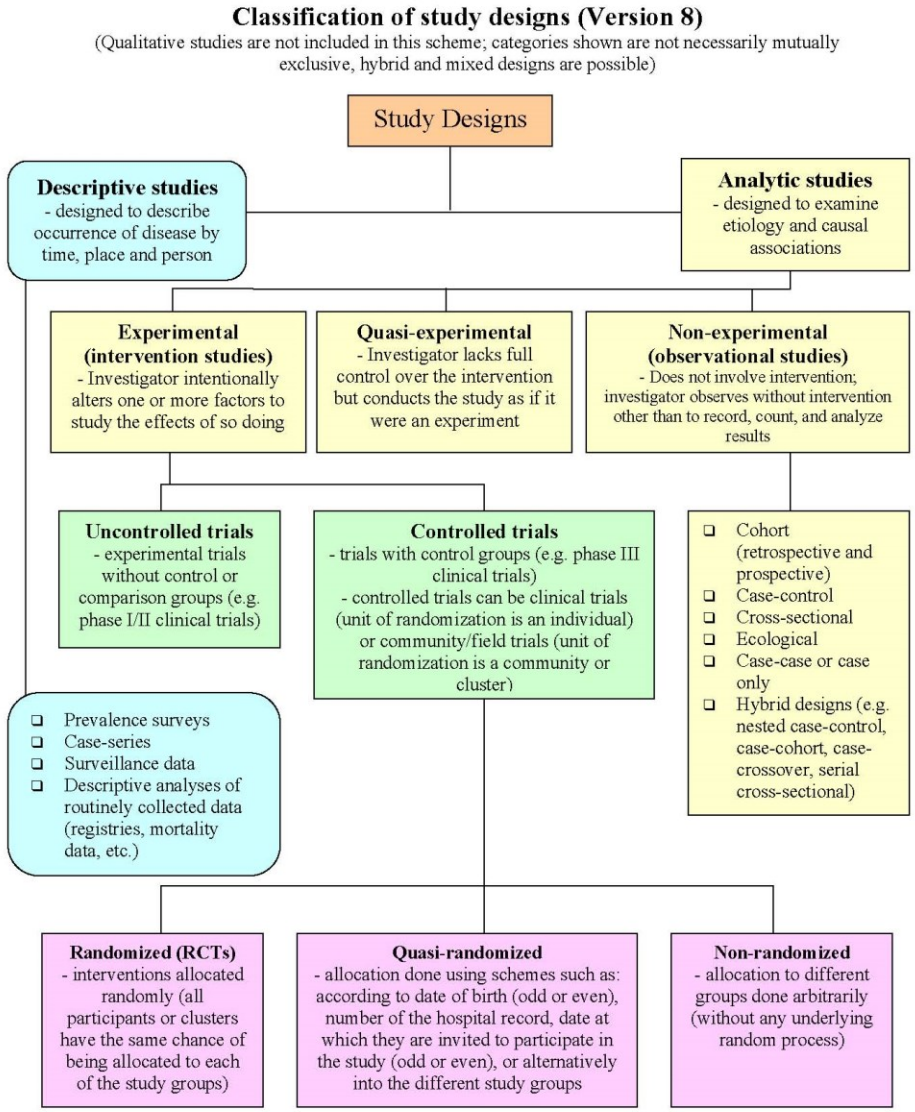
Figure 1. Results From Sensitivity Analyses Dividing Trials in High- and Low-Quality Strata, Using 25 Different Quality Assessment Scales



Relative risks (RRs) for deep vein thrombosis with 95% confidence intervals (CIs) are shown. LMWH indicates low-molecular-weight heparin. Black squares indicate estimates from high-quality trials and open squares indicate estimates from low-quality trials. Arrows indicate that the values are outside the range of the x axis. Broken line indicates combined estimate from all 17 trials. Solid line indicates null effect line. The scales are arranged in decreasing order of the RRs in trials deemed to be of high quality. Asterisk indicates unpublished scale.

Juni et al.
JAMA 1999

Quality instruments will depend on study design



Checklists and scales for RCTs

Quality assessment

- Quality assessment of RCTs
  - Users' Guides to the Medical Literature checklist:
    - Were patients randomized?
    - Was randomization (allocation) concealed?
    - Were patients analyzed in the groups to which they were randomized? (intention-to-treat analysis)
    - Were patients in the treatment and control groups similar with respect to known prognostic variables?
    - Were patients aware of group allocation?
    - Were clinicians aware of group allocation?
    - Were outcome assessors aware of group allocation?
    - Was follow-up complete?

Guyatt et al. Users' Guides to the Medical Literature. Chicago: AMA Press, 2002

Quality assessment: Jadad Scale for quality of RCTs

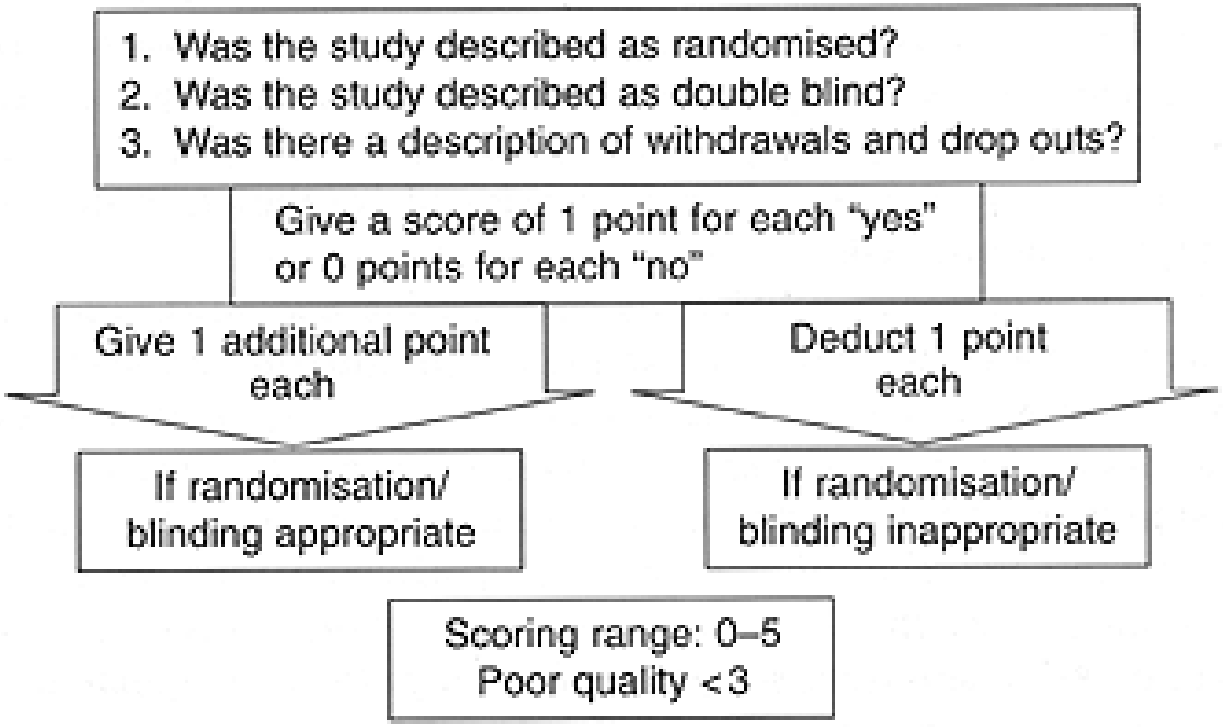


Figure 4.1 Validated quality scale. (From Jadad et al.[1])

"The use of this scale is explicitly discouraged. As well as suffering from the generic problems of scales, it has a strong emphasis on reporting rather than conduct, and does not cover one of the most important potential biases in randomized trials, namely allocation concealment"

- Cochrane Handbook

Jadad AR, et al. Assessing the quality of reports on randomized clinical trials: Is blinding necessary? *Controlled Clin Trials* 1996;**17**:1-12. URL: http://www.bmjpg.com/rct/chapter4.html

Cochrane tool for risk of bias in RCTs

Table 8.5.a: The Cochrane Collaboration's tool for assessing risk of bias

Domain	Description	Review authors' judgement
Sequence generation.	Describe the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.	Was the allocation sequence adequately generated?
Allocation concealment.	Describe the method used to conceal the allocation sequence in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.	Was allocation adequately concealed?
Blinding of participants, personnel and outcome assessors *Assessments should be made for each main outcome (or class of outcomes).*	Describe all measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received. Provide any information relating to whether the intended blinding was effective.	Was knowledge of the allocated intervention adequately prevented during the study?
Incomplete outcome data *Assessments should be made for each main outcome (or class of outcomes).*	Describe the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis. State whether attrition and exclusions were reported, the numbers in each intervention group (compared with total randomized participants), reasons for attrition/exclusions where reported, and any re-inclusions in analyses performed by the review authors.	Were incomplete outcome data adequately addressed?
Selective outcome reporting.	State how the possibility of selective outcome reporting was examined by the review authors, and what was found.	Are reports of the study free of suggestion of selective outcome reporting?
Other sources of bias.	State any important concerns about bias not addressed in the other domains in the tool. If particular questions/entries were pre-specified in the review's protocol, responses should be provided for each question/entry.	Was the study apparently free of other problems that could put it at a high risk of bias?

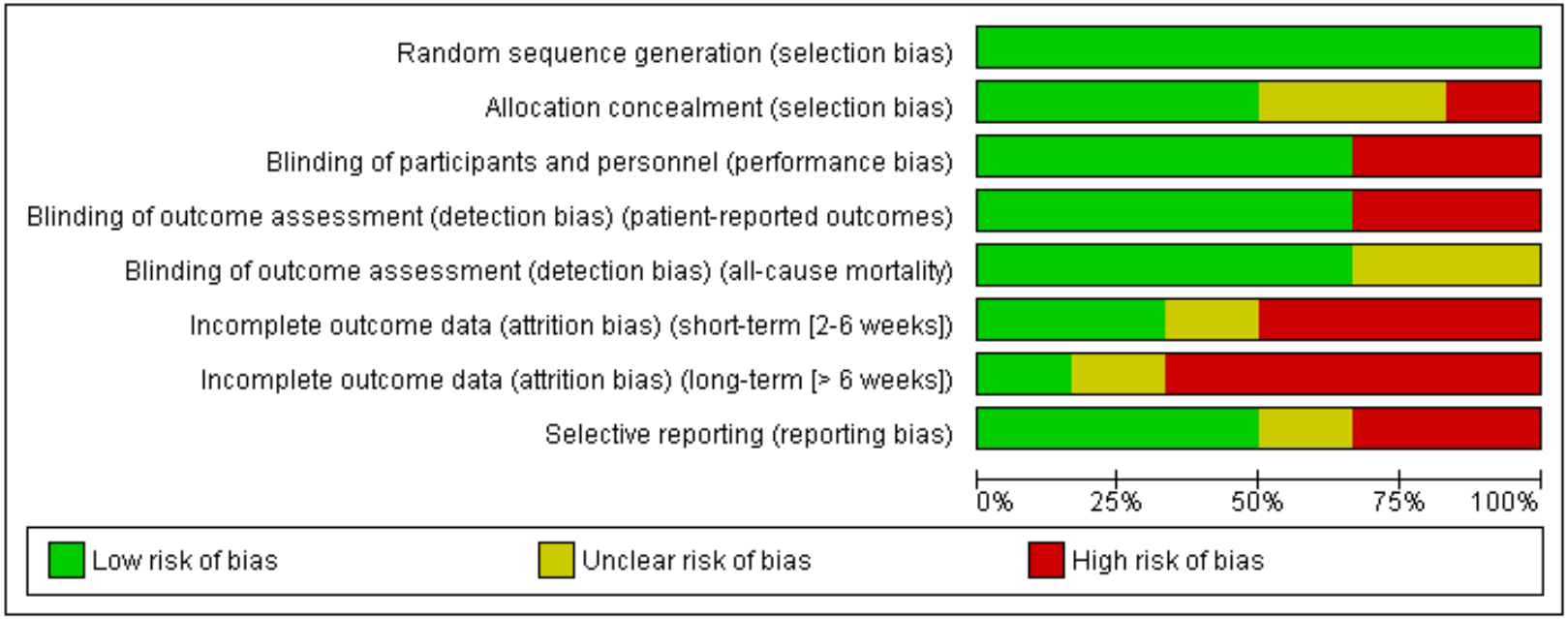
http://www.cochrane-handbook.org/

Figure 8.6.a: Example of a 'Risk of bias' table for a single study (fictional)

Entry	Judgement	Support for judgement
Random sequence generation (selection bias)	Low risk.	Quote: "patients were randomly allocated." Comment: Probably done, since earlier reports from the same investigators clearly describe use of random sequences (Cartwright 1980).
Allocation concealment (selection bias)	High risk.	Quote: "...using a table of random numbers." Comment: Probably not done.
Blinding of participants and personnel (performance bias)	Low risk.	Quote: "double blind, double dummy"; "High and low dose tablets or capsules were indistinguishable in all aspects of their outward appearance. For each drug an identically matched placebo was available (the success of blinding was evaluated by examining the drugs before distribution)." Comment: Probably done.
Blinding of outcome assessment (detection bias) (patient-reported outcomes)	Low risk.	Quote: "double blind". Comment: Probably done.
Blinding of outcome assessment (detection bias) (Mortality)	Low risk.	Obtained from medical records; review authors do not believe this will introduce bias.
Incomplete outcome data addressed (attrition bias) (Short-term outcomes (2-6 weeks))	High risk.	4 weeks: 17/110 missing from intervention group (9 due to 'lack of efficacy'); 7/113 missing from control group (2 due to 'lack of efficacy').
Incomplete outcome data addressed (attrition bias) (Longer-term outcomes (>6 weeks))	High risk.	12 weeks: 31/110 missing from intervention group; 18/113 missing from control group. Reasons differ across groups.
Selective reporting (reporting bias)	High risk.	Three rating scales for cognition listed in Methods, but only one (with statistically significant results) is reported.

http://www.cochrane-handbook.org/

Figure 8.6.b: Example of a 'Risk of bias graph' figure



http://www.cochrane-handbook.org/

Table 8.7.a: Possible approach for *summary assessments* of the risk of bias for each important outcome (across domains) within and across studies

Risk of bias	Interpretation	Within a study	Across studies
Low risk of bias.	Plausible bias unlikely to seriously alter the results.	Low risk of bias for all key domains.	Most information is from studies at low risk of bias.
Unclear risk of bias.	Plausible bias that raises some doubt about the results.	Unclear risk of bias for one or more key domains.	Most information is from studies at low or unclear risk of bias.
High risk of bias.	Plausible bias that seriously weakens confidence in the results.	High risk of bias for one or more key domains.	The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results.

Trials

BioMed Central

Methodology

Open Access

Randomized trials published in some Chinese journals: how many are randomized?

Taixiang Wu*[1], Youping Li[1], Zhaoxiang Bian[2], Guanjian Liu[1] and David Moher[3]

Results: From an initial sample of 37,313 articles identified in the China National Knowledge Infrastructure database, we found 3137 apparent randomized controlled trials. Of these, 1452 were studies of conventional medicine (published in 411 journals) and 1685 were studies of traditional Chinese medicine (published in 352 journals). Interviews with the authors of 2235 of these reports revealed that only 207 studies adhered to accepted methodology for randomization and could on those grounds be deemed authentic randomized controlled trials (6.8%, 95% confidence interval 5.9–7.7). There was no statistically significant difference in the rate of authenticity between randomized controlled trials of traditional interventions and those of conventional interventions. Randomized controlled trials conducted at hospitals affiliated to medical universities were more likely to be authentic than trials conducted at level 3 and level 2 hospitals (relative risk 1.58, 95% confidence interval 1.18–2.13, and relative risk 14.42, 95% confidence interval 9.40–22.10, respectively). The likelihood of authenticity was higher in level 3 hospitals than in level 2 hospitals (relative risk 9.32, 95% confidence interval 5.83–14.89). All randomized controlled trials of pre-market drug clinical trial were authentic by our criteria. Of the trials conducted at university-affiliated hospitals, 56.3% were authentic (95% confidence interval 32.0–81.0).

Conclusion: Most reports of randomized controlled trials published in some Chinese journals lacked an adequate description of randomization. Similarly, most so called 'randomized controlled trials' were not real randomized controlled trials owing toa lack of adequate understanding on the part of the authors of rigorous clinical trial design. All randomized controlled trials of pre-market drug clinical trial included in this research were authentic. Randomized controlled trials conducted by authors in high level hospitals, especially in hospitals affiliated to medical universities had a higher rate of authenticity. That so many non-randomized controlled trials were published as randomized controlled trials reflected the fact that peer review needs to be improved and a good practice guide for peer review including how to identify the authenticity of the study urgently needs to be developed.

ELSEVIER

Journal of Clinical Epidemiology 55 (2002) 787–790

Epidemiology

In the dark
The reporting of blinding status in randomized controlled trials

Victor M. Montori[a,*], Mohit Bhandari[b], P.J. Devereaux[c], Braden J. Manns[d], William A. Ghali[d], Gordon H. Guyatt[e]

[a]*Department of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA*
[b]*Department of Surgery, McMaster University, Hamilton, ON, Canada*
[c]*Department of Medicine, McMaster University, Hamilton, Ontario, Canada*
[d]*Department of Medicine and Community Health Sciences, University of Calgary, Alberta, Canada*
[e]*Departments of Medicine and Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada*

Received 10 August 2001; received in revised form 11 April 2002; accepted 26 April 2002

Abstract

To determine the quality of reporting of blinding in randomized controlled trials (RCTs), we evaluated 40 consecutive RCTs published in each of five leading journals. We noted whether authors reported the blinding status of participants, health care providers, data collectors, judicial assessors of outcomes, data analysts, and manuscript writers. Explicit reporting of blinding status occurred in <25% of RCTs for all groups. Eighty-three RCTs, reported as double-blind, provided eight combinations of blinded groups. In conclusion, prestigious journals do not currently report blinding status optimally. To do so, journals should abandon the term "double blind" and explicitly report the blinding status of the groups involved in RCTs. Until such reporting occurs, clinicians will be left with uncertainty about the validity of RCTs that guide their clinical practice. © 2002 Elsevier Science Inc. All rights reserved.

Checklists and scales for observational studies (case-control and cohort)

Quality assessment

- Quality assessment of observational studies: cohort studies
  - The Newcastle-Ottawa Scale
    - **Selection**
      - 1) Representativeness of the exposed cohort
      - 2) Selection of the non exposed cohort
      - 3) Ascertainment of exposure
      - 4) Demonstration that outcome of interest was not present at start of study
    - **Comparability**
      - 1) Comparability of cohorts on the basis of the design or analysis
    - **Outcome**
      - 1) Assessment of outcome
      - 2) Was follow-up long enough for outcomes to occur
      - 3) Adequacy of follow up of cohorts

http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm

Quality assessment

- Quality assessment of observational studies: case-control studies:
  - The Newcastle-Ottawa Scale:
    - **Selection**
      - 1) Is the case definition adequate?
      - 2) Representativeness of the cases
      - 3) Selection of Controls
      - 4) Definition of Controls
    - **Comparability**
      - 1) Comparability of cases and controls
    - **Exposure**
      - 1) Ascertainment of exposure
      - 2) Same method of ascertainment for cases and controls
      - 3) Non-Response rate

http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm

Methodology Checklist 3: Cohort studies

SIGN

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper really a cohort study? If in doubt, check the study design algorithm available from SIGN and make sure you have the correct checklist.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist..

Reason for rejection: 1. Paper not relevant to key question ☐ 2. Other reason ☐ (please specify):

Checklist completed by:

SECTION 1: INTERNAL VALIDITY

	In a well conducted cohort study:	*In this study the criterion is:*
1.1	The study addresses an appropriate and clearly focused question.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
	SELECTION OF SUBJECTS	
1.2	The two groups being studied are selected from source populations that are comparable in all respects other than the factor under investigation.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.3	The study indicates how many of the people asked to take part did so, in each of the groups being studied.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.4	The likelihood that some eligible subjects might have the outcome at the time of enrolment is assessed and taken into account in the analysis.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.5	What percentage of individuals or clusters recruited into each arm of the study dropped out before the study was completed.	
1.6	Comparison is made between full participants and those lost to follow up, by exposure status.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
	ASSESSMENT	
1.7	The outcomes are clearly defined.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.8	The assessment of outcome is made blind to exposure status.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.9	Where blinding was not possible, there is some recognition that knowledge of exposure status could have influenced the assessment of outcome.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.10	The measure of assessment of exposure is reliable.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.11	Evidence from other sources is used to demonstrate that the method of outcome assessment is valid and reliable.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
1.12	Exposure level or prognostic factor is assessed more than once.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
	CONFOUNDING	
1.13	The main potential confounders are identified and taken into account in the design and analysis.	Well covered / Adequately addressed / Poorly addressed / Not addressed / Not reported / Not applicable
	STATISTICAL ANALYSIS	
1.14	Have confidence intervals been provided?	

SECTION 2: OVERALL ASSESSMENT OF THE STUDY

2.1	How well was the study done to minimise the risk of bias or confounding, and to establish a causal relationship between exposure and effect? *Code ++, +, or –*	
2.2	Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, are you certain that the overall effect is due to the study intervention?	
2.3	Are the results of this study directly applicable to the patient group targeted in this guideline?	
2.4	**Notes.** Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question.	

The following section is provided for non-SIGN users of this checklist and is being developed to conform to the standards set by the Guidelines International Network Evidence Tables Working Group.

Members of SIGN guideline groups do not need to complete this section.

SECTION 3: DESCRIPTION OF THE STUDY

PLEASE PRINT CLEARLY

3.1	*Do we know who the study was funded by?*	☐ Academic Institution ☐ Healthcare Industry ☐ Government ☐ NGO ☐ Public funds ☐ Other

© Scottish Intercollegiate Guidelines Network, March 2004

© Scottish Intercollegiate Guidelines Network, March 2004

http://www.sign.ac.uk/methodology/checklists.html

Methodology Checklist 4: Case-control studies

SIGN

Study identification (*Include author, title, year of publication, journal title, pages*)

Guideline topic:	Key Question No:

Before completing this checklist, consider:

1. Is the paper really a case-control study? If in doubt, check the study design algorithm available from SIGN and make sure you have the correct checklist.
2. Is the paper relevant to key question? Analyse using PICO (Patient or Population Intervention Comparison Outcome). IF NO REJECT (give reason below). IF YES complete the checklist.

Reason for rejection: Reason for rejection: 1. Paper not relevant to key question ☐ 2. Other reason ☐ (please specify):

Checklist completed by:

SECTION 1: INTERNAL VALIDITY

	In an well conducted case control study:	***In this study the criterion is:***	
1.1	The study addresses an appropriate and clearly focused question	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
SELECTION OF SUBJECTS			
1.2	The cases and controls are taken from comparable populations	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.3	The same exclusion criteria are used for both cases and controls	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.4	What percentage of each group (cases and controls) participated in the study?	Cases: Controls:	
1.5	Comparison is made between participants and non-participants to establish their similarities or differences	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.6	Cases are clearly defined and differentiated from controls	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.7	It is clearly established that controls are non-cases	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
ASSESSMENT			
1.8	Measures will have been taken to prevent knowledge of primary exposure influencing case ascertainment	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
1.9	Exposure status is measured in a standard, valid and reliable way	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
CONFOUNDING			
1.10	The main potential confounders are identified and taken into account in the design and analysis	Well covered Adequately addressed Poorly addressed	Not addressed Not reported Not applicable
STATISTICAL ANALYSIS			
1.11	Confidence intervals are provided		

SECTION 2: OVERALL ASSESSMENT OF THE STUDY

2.1	How well was the study done to minimise the risk of bias or confounding? *Code ++, +, or –*	
2.2	Taking into account clinical considerations, your evaluation of the methodology used, and the statistical power of the study, are you certain that the overall effect is due to the study intervention?	
2.3	Are the results of this study directly applicable to the patient group targeted by this guideline?	
2.4	**Notes.** Summarise the authors conclusions. Add any comments on your own assessment of the study, and the extent to which it answers your question.	

The following section is provided for non-SIGN users of this checklist and is being developed to conform to the standards set by the Guidelines International Network Evidence Tables Working Group.

Members of SIGN guideline groups do not need to complete this section.

SECTION 3: DESCRIPTION OF THE STUDY

PLEASE PRINT CLEARLY

3.1	*Do we know who the study was funded by?*	☐ Academic Institution ☐ Healthcare Industry ☐ Government ☐ NGO ☐ Public funds ☐ Other
3.2	*How many centres are patients recruited from?*	
3.3	*From which countries are patients selected? (Select all those involved. Note additional countries after "Other")*	☐ Scotland ☐ UK ☐ USA ☐ Canada ☐ Australia ☐ New Zealand ☐ France ☐ Germany ☐ Italy ☐ Netherlands ☐ Scandinavia ☐ Spain ☐ Other:

http://www.sign.ac.uk/methodology/checklists.html

Special Issue Paper

Research Synthesis Methods

Received 23 January 2013, Accepted 3 February 2013 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/jrsm.1077

Checklists of methodological issues for review authors to consider when including non-randomized studies in systematic reviews

George A Wells,[a,b*†] Beverley Shea,[c] Julian PT Higgins,[d,e] Jonathan Sterne,[f] Peter Tugwell[b,g] and Barnaby C Reeves[h]

Checklists for diagnostic studies

QUADAS tool for quality assessment of diagnostic studies

Table 1: QUADAS

Item #	Description
1.	Was the spectrum of patients representative of the patients who will receive the test in practice?
2.	Were selection criteria clearly described?
3.	Is the reference standard likely to correctly classify the target condition?
4.	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? (disease progression bias)
5.	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis? (partial verification bias)
6.	Did patients receive the same reference standard regardless of the index test result? (differential verification bias)
7.	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)? (incorporation bias)
8.	Was the execution of the index test described in sufficient detail to permit replication of the test?
9.	Was the execution of the reference standard described in sufficient detail to permit its replication?
10.	Were the index test results interpreted without knowledge of the results of the reference standard? (test review bias)
11.	Were the reference standard results interpreted without knowledge of the results of the index test? (diagnostic review bias)
12.	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (clinical review bias)
13.	Were uninterpretable/ intermediate test results reported?
14.	Were withdrawals from the study explained?

Whiting P et al. BMC Med Res Meth 2003

Annals of Internal Medicine | Research and Reporting Methods

QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies

Penny F. Whiting, PhD; Anne W.S. Rutjes, PhD; Marie E. Westwood, PhD; Susan Mallett, PhD; Jonathan J. Deeks, PhD; Johannes B. Reitsma, MD, PhD; Mariska M.G. Leeflang, PhD; Jonathan A.C. Sterne, PhD; Patrick M.M. Bossuyt, PhD; and the QUADAS-2 Group*

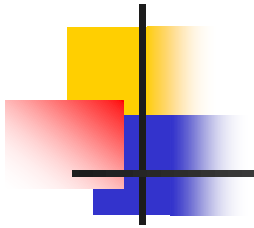
QUADAS-2

Phase 1: State the review question:

Patients (setting, intended use of index test, presentation, prior testing):
Index test(s):
Reference standard and target condition:

Phase 2: Draw a flow diagram for the primary study

http://www.bris.ac.uk/quadas/quadas-2



For sample quality assessment checklists, check the USB drive!

Study quality vs reporting

- Methodological quality versus quality of reporting
  - It is hard to separate quality of reporting from methodological quality
  - "Not reported" is not always "not done"
  - It is best to have categories such as:
    - Criteria met
    - Criteria not met
    - Not reported or can't tell
  - Reviewers often will have to write to the authors of primary studies and obtain missing information
    - Emailing authors might be a good strategy

Study quality vs. study reporting

Data from a meta-analysis of NAAT for TB meningitis (Pai et al. *Lancet Infect Dis* 2003)

Characteristic	Before contact % [N = 49]	After contact % [N = 49]
Blinding		
Double blind	12	35
Single blind	14	24
Unblinded	0	10
Not reported	74	31
Sampling		
Consecutive/random	18	49
Not consecutive/random	6	20
Not reported	76	31
Data collection		
Prospective	51	61
Retrospective	0	4
Both	2	10
Not reported	47	25

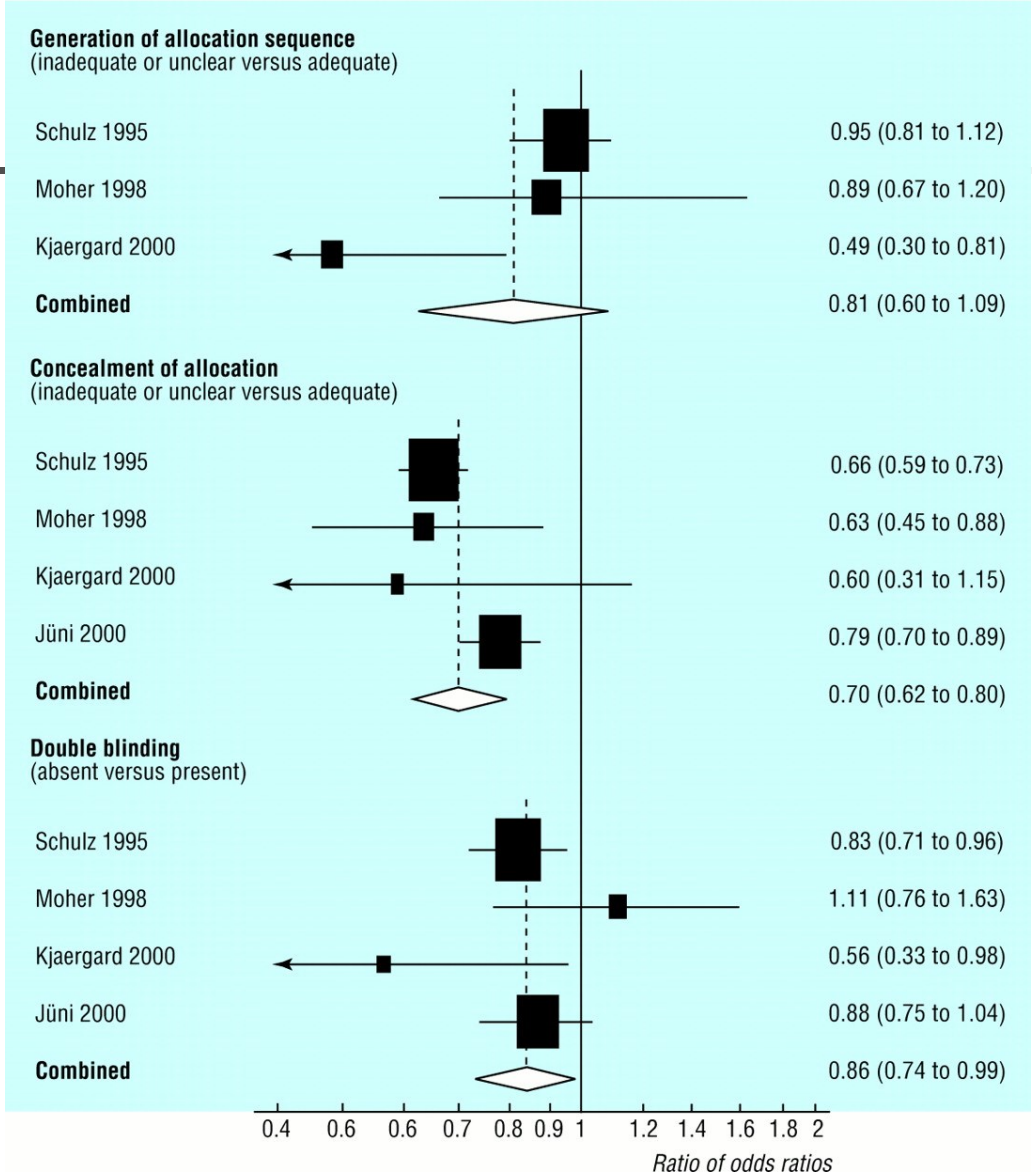
Pai M, et al. Quality assessment in meta-analyses of diagnostic studies: what difference does email contact with authors make? Abstract: XI Cochrane Colloquium, Barcelona, Spain, 26 – 31, October 2003.

Initiatives to improve quality of reporting

- CONSORT: reporting of RCTs
- STARD: reporting of diagnostic studies
- STROBE: reporting of observational studies
- PRISMA: reporting of meta-analyses of RCTs
- MOOSE: reporting of meta-analyses of observational studies

http://www.equator-network.org/

Do design flaws affect RCT results?



Juni, P. et al. BMJ 2001;323:42-46

Do design flaws affect diagnostic study results?

487 diagnostic studies

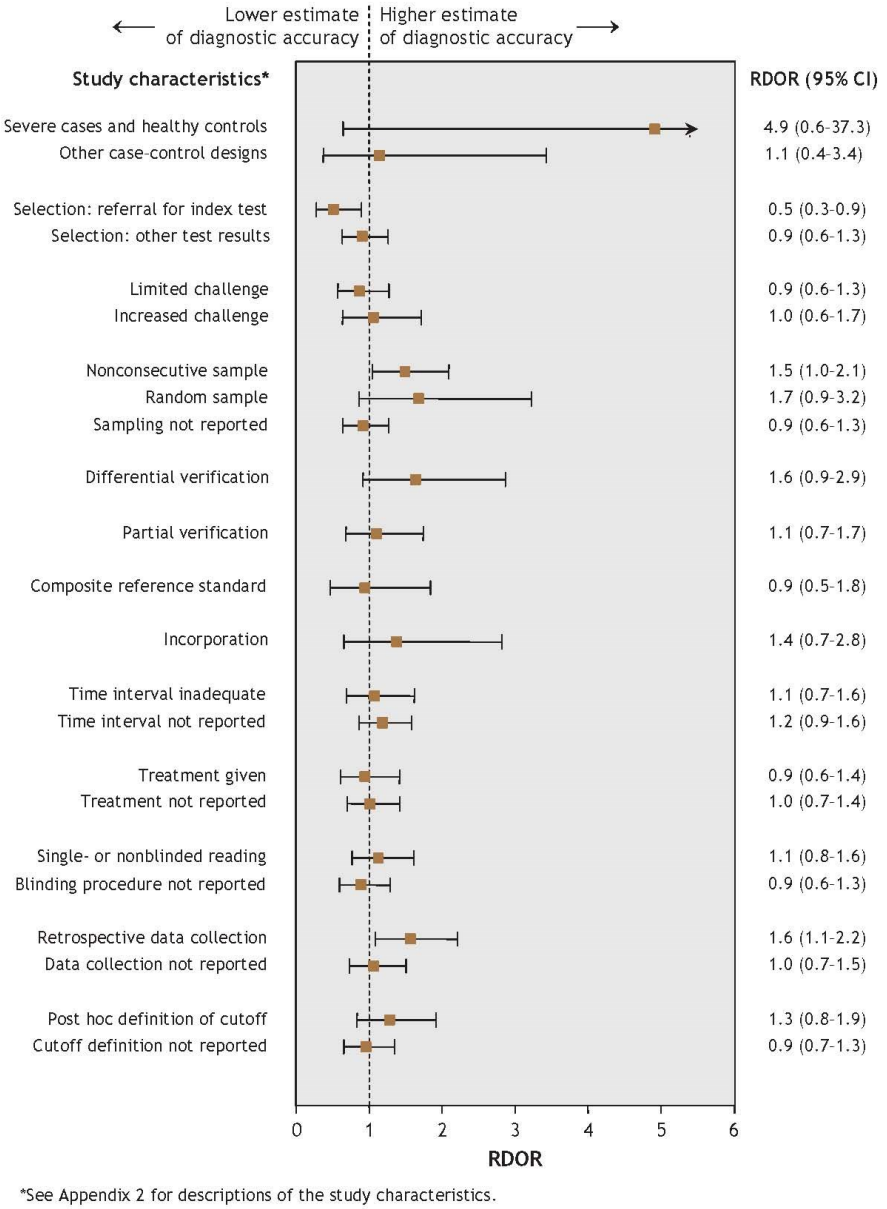


Fig. 2: Effects of study design characteristics on estimates of diagnostic accuracy. RDOR = relative diagnostic odds ratio (adjusted RDORs were estimated in a multivariable random-effects meta-epidemiologic regression model).

Rutjes et al. CMAJ 2006

Quality assessment

- How can one use quality information during analysis?
  - Narrative discussion of impact of quality on results
  - Display study quality and results in a tabular format
  - Weight the data by quality (not recommended)
  - Subgroup analysis by quality
  - Include quality as a covariate in meta-regression

Tabular display of quality information

REVIEW

Male circumcision and risk of syphilis, chancroid, and genital herpes: a systematic review and meta-analysis

H A Weiss, S L Thomas, S K Munabi, R J Hayes

Sex Transm Infect 2006;82:101–110. doi: 10.1136/sti.2005.017442

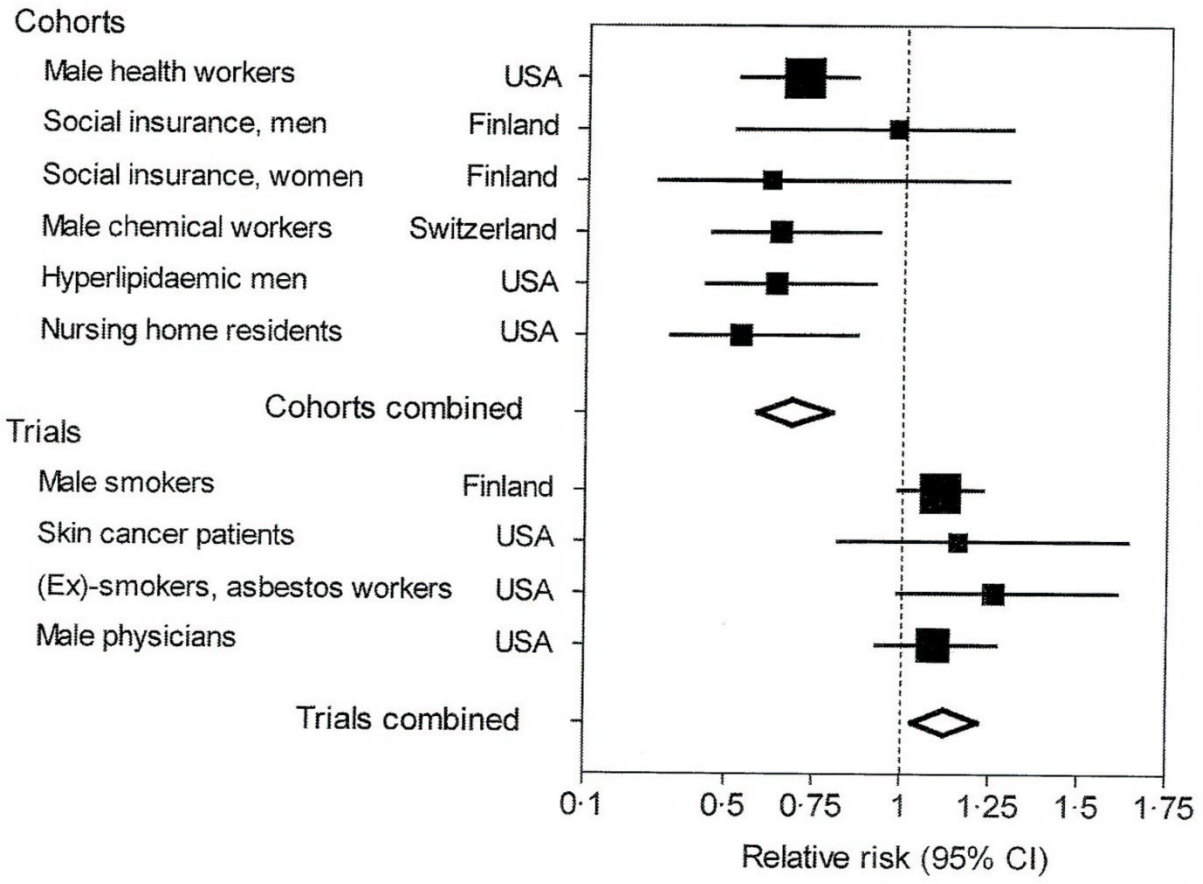
Table 1 Summary of studies of the association of male circumcision and HSV-2 serostatus

First author	Design	Location, date of study	Study population	Study size	HSV-2	Circumcised	Assessment of circumcision	Crude RR* (95% CI)	Adjusted RR* (95% CI)	Adjusted for
Auvert[48]	Cross sectional	Carltonville, South Africa 1999	Young adults aged 14–24 years	676	15%	3%	Self report	1.20 (0.48 to 2.96)	–	
Gottlieb[24]	Cohort	Five cities, USA 1993–6	STD clinic	1120	9%	71%†	Clinical examination	0.88 (0.5 to 1.4)	1.0 (0.6–1.6)	Age, race, city, HSV1 status, condom use with occasional partner
Gray[25]‡	Nested case-control	Rakai, Uganda 1994–8	General	674	70%	18%†	Self report	0.82 (0.68 to 0.99)	0.81 (0.67 to 0.97)	Age, marital status, condom use, number of lifetime partners
Kapiga[47]	Cross sectional	Moshi, Tanzania 2000	Bar workers	206	29%	95%	Clinical examination	1.07 (0.40 to 2.88)	0.56 (0.13 to 2.5)	Age only
Lavreys[43]	Cross sectional	Mombasa, Kenya 1993–7	Trucking employees	113	46%	57%	Clinical examination	1.18 (0.78 to 1.79)	to	
Obasi[44]	Cross sectional	Mwanza, Tanzania 1992–3	General	133	23%	23%	Self report	0.68 (0.28 to 1.62)	0.39 (0.1 to 1.52)	Age, residence, mobility, marital status, lifetime partners, TPHA status
Reynolds[26]	Cohort	Pune, India 1993–2000	STD clinic attenders	2298	14%	8%†	Clinical examination	0.89 (0.48 to 1.53)	0.91 (0.51 to 1.64)	Age, religion, education, living with family, year, marital status, number of sex partners, number of female sex worker partners, condom use, tattoos, medical injections
Suligoi[46]	Cross sectional	Garoua, Cameroon 1997–8	Outpatients	82	24%	91%†	Self report	0.84 (0.24 to 2.90)	–	
Weiss[45]	Cross sectional	Kisumu, Kenya 1997	General	583	35%	27%†	Clinical examination	0.65 (0.47 to 0.90)	0.73 (0.47 to 1.13)	Age, marital status, ethnic group and number of lifetime partners.
Weiss[45]	Cross sectional	Ndola, Zambia 1997	General	607	36%	9%†	Clinical examination	1.20 (0.81 to 1.77)	1.04 (0.74 to 1.44)	Age, marital status and number of lifetime partners.

*Rate ratio in references 24, 26. Odds ratio in references 45, 47; Prevalence ratio in references 25, 44, 45, 46, 48.
†Circumcision before sexual debut.
‡Nested HIV incident case-control study of HIV seroconvertors and matched seronegative controls.

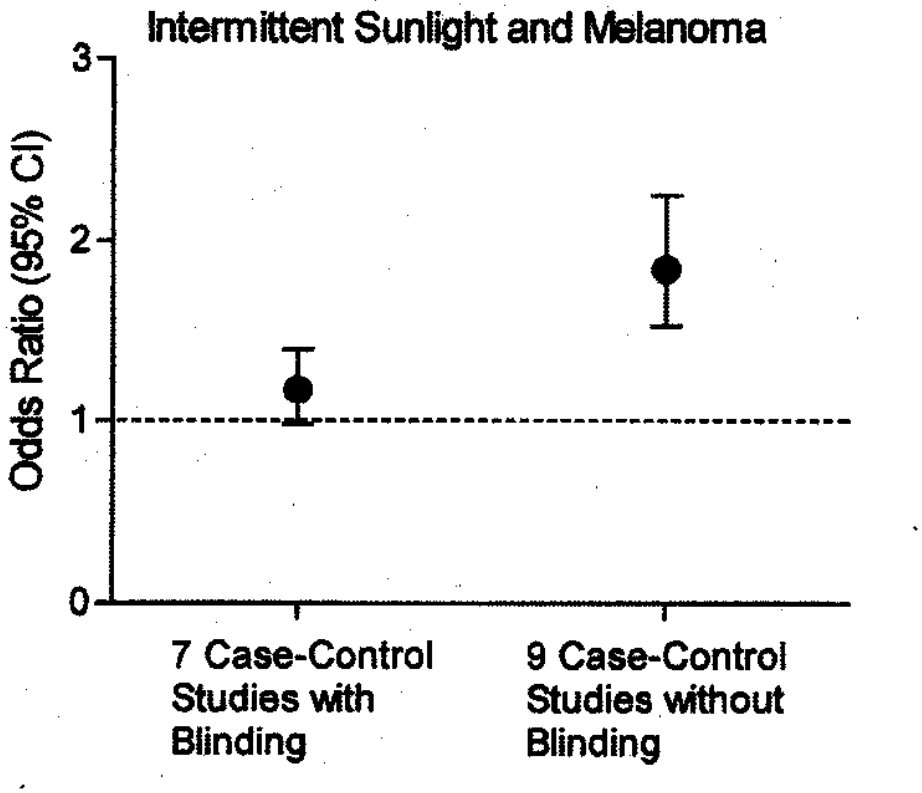
Impact of quality on results: example 1

Beta-carotene intake and cardiovascular mortality



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

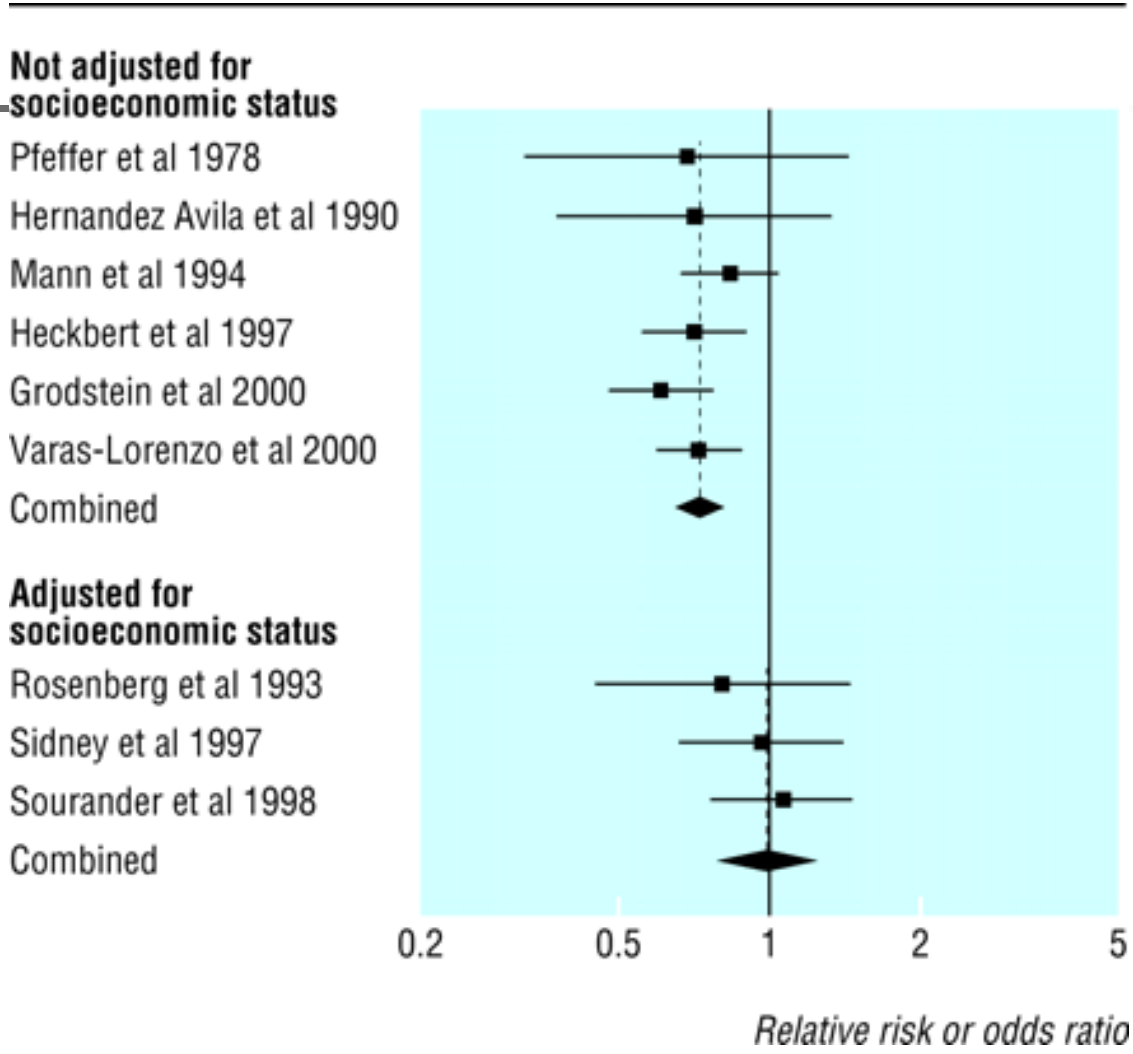
Impact of quality on results: example 2



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

Impact of quality on results: example 3

Hormone replacement therapy and cardiovascular disease



BMJ 2004;329:868-869 (16 October)

Impact of quality on results: example 4

Coffee consumption and risk of coronary heart disease: A meta-analysis

Francesco Sofi [a,d,*], Andrea A. Conti [a,b], Anna Maria Gori [a,d], Maria Luisa Eliana Luisi [b], Alessandro Casini [c,d], Rosanna Abbate [a,d], Gian Franco Gensini [a,b]

Abstract *Background and aims:* During the past three decades the relationship between habitual coffee drinking and coronary heart disease (CHD) has been assessed in numerous studies, with conflicting results. The aim of this study was to systematically examine the data published on the association between habitual coffee consumption and risk of CHD.
Methods and results: Thirteen case–control and 10 cohort studies were included. Case–control studies incorporated 9487 cases of CHD and 27,747 controls, and cohort studies included a total of 403,631 participants that were followed for between 3 and 44 years. The summary of odds ratios (OR) for the case–control studies showed statistically significant associations between coffee consumption and CHD for the highest intake group (>4 cups/day), OR 1.83 (95% CI 1.49–2.24; $P < 0.0001$), and for the second highest category (3–4 cups/day), OR 1.33 (95% CI 1.04–1.71; $P < 0.0001$), while no significant association emerged for low daily coffee intake (≤2 cups/day), OR 1.03 (95% CI 0.87–1.21; $P = 0.45$). The analysis of long-term follow-up cohort studies did not show any association between the consumption of coffee and CHD, with a relative risk (RR) of 1.16 (95% CI 0.95–1.41; $P = 0.14$) for the highest category, and 1.05 (95% CI 0.90–1.22; $P = 0.57$) and

Case-control:
OR = 1.83

Cohort:
RR = 1.16

Nutr Metab Cardiovasc Dis 2007

Impact of quality on results: example 5

Meta-Analysis of Mobile Phone Use and Risk of Tumor

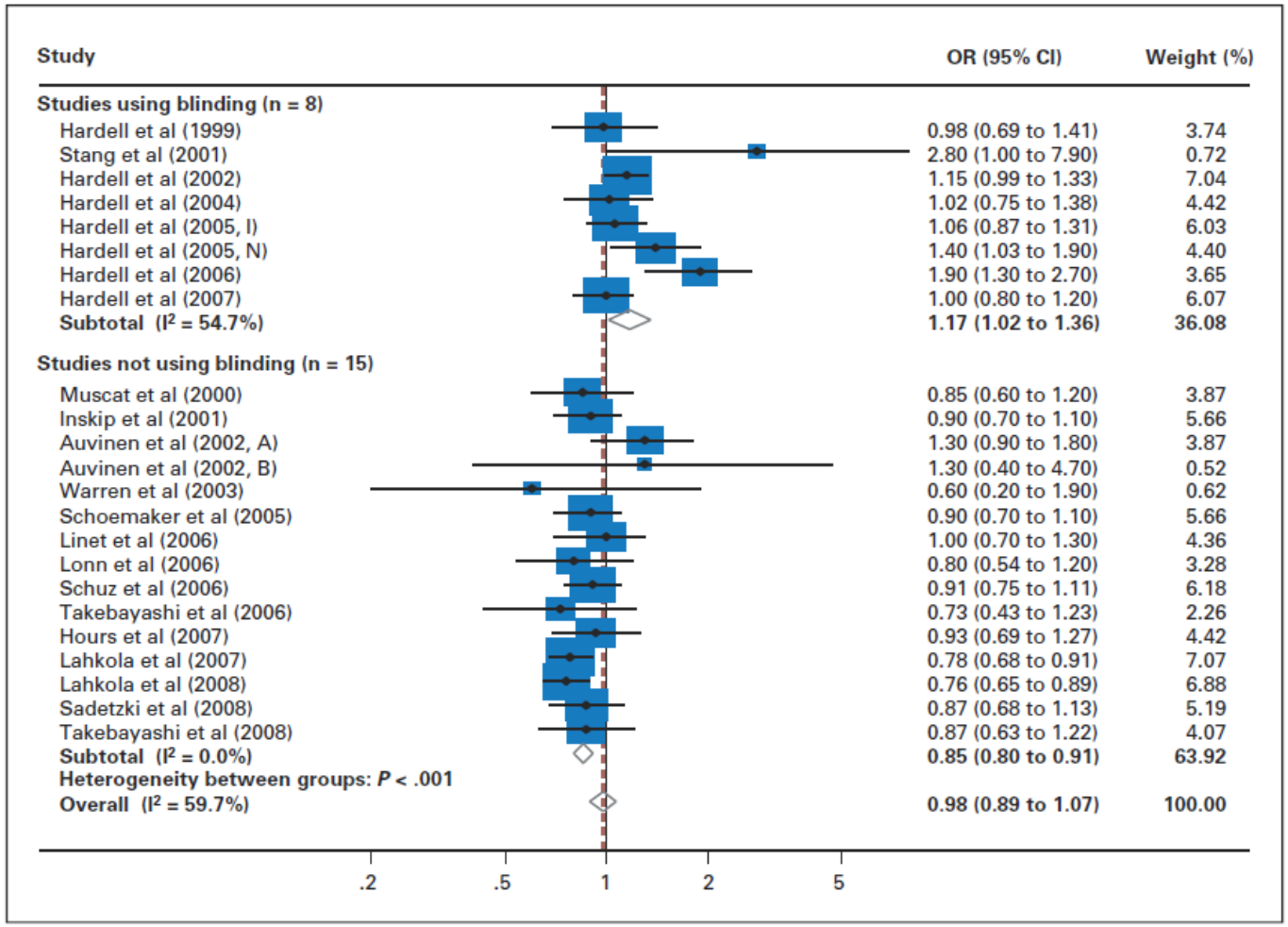


Fig 2. Overall use of mobile phones and the risk of tumors in a random-effects model meta-analysis of case-control studies[7-28] by the use of blinding at an interview for exposure measurements (n = 23). OR, odds ratio; Hardell et al (2005, I) indicates reference 15; Hardell et al (2005, N) indicates reference 16.

Myung et al. J Clin Oncol 2009

Data extraction

- Extraction of data and quality assessment are usually done together
- Development of a clear, well-designed data extraction form and coding instruction manual is crucial
- Pilot testing is absolutely necessary!
- Reproducibility reduces error and subjectivity
- Contacting authors for missing data increases the accuracy of collected data

Commercial serological tests for the diagnosis of tuberculosis: an updated systematic review and meta-analysis, Steingart et al PLoS Medicine (in press)

Commercial Serological Antibody Detection Test- Data Extraction 2010 – Pulmonary TB

ID	
IDsub	
Author	
Title	
Year	
Language	
Industry involvement	0-No 1-Yes 9-NR/unclear
If yes, check all that apply	If Yes, characterize type (Select one: answers ordered from least to most industry involvement) __Donation of test materials or kits __Receipt of educational support, grants, or speaking fees __Work/financial relationship (author is an employee/consultant or owns company stock) __Involvement in design, analysis, or manuscript production
PTB, reference standard	1-Culture 2-Smear 6-Culture &/or Smear 7-Bact Confirmation
If culture used as ref standard, was culture	1-Solid Culture 2-Liquid Culture 3-Both Solid and Liquid Culture 4-Culture, type NR
PTB smear status	1- Positive 2-Negative 9-NR/Unclear
Setting	1-Outpatient 2-Inpatient 3-Both out- and inpatient 4-Lab-based 5-Other, specify 9-NR/Unclear

Does Bleach Processing Increase the Accuracy of Sputum Smear Microscopy for Diagnosing Pulmonary Tuberculosis? Cattamanchi . J Clin Microbiol. 2010

PrimKey: New
Year:
ID:
Publication Type:
SubID:
Country:
Author:

Design:
StudyPopulation:
Data collection:
Health Care Setting:
Patients or specimens?:
Patient Selection:
Age Cateogory:
HIV-positives included?:
Specify % HIV positive:
Eligibility criteria clearly described?:
Same reference standard for DS and PS?:
PS method described in sufficient detail?:
Smears interpreted w/o knowing cx result?:
DS and CS interpreted blindly?:

Number of specimens collected per patient:
How were sputa collected?:
How were smears prepared?:
Was DM performed in a field lab?:
Was PM performed in a field lab?:
Were DM and PM done in same lab?:
Was microscopist training reported?:
Was an EQA program in place?:

Comments:

Chemical Processing Method:
Physical Processing Method:
Type of Stain Used:
Smear positive definition (>= # AFB/100 hpf:
Gold standard used?:
Gold Standard Type:

Centrifugation speed: rpm or g?:
Centrifugation time (minutes):
Sedimentation time (hours):

Results - All Patients

Gold Standard Used				No Gold Standard Used	
Direct Smear				Direct Smear	
TP		FP		Number Positive	
FN		TN		Total	
Processed Smear				Processed Smear	
TP		FP		Number Positive	
FN		TN		Total	

Results - HIV-infected Patients

Gold Standard Used				No Gold Standard Used	
Direct Smear				Direct Smear	
TP		FP		Number Positive	
FN		TN		Total:	
Processed Smear				Processed Smear	
TP		FP		Number Positive	
FN		TN		Total:	

Cochrane SR on Xpert MTB/RF by Steingart K et al.

Study Identification									2 x 2 data										
									MTB_ALL										
ID	sub	author	YEAR	StudyP	REF_STD_MTB	REF_STD_P	SETTING	COUNTRY	N_ALL	N_TH F_I	TP	FP	FN	TN	IND	TOT	SENS	SPEC	IND	TB Prev
1	a	Boehme	2010	BOTH	LJ AND MGIT960	LJ_PM	Azerbaijan	Azerbaijan	218	49	141	1	8	68	2	218	94.6%	98.6%	0.9%	0.68
									216		123	1	24	68	4	216	83.7%	98.6%	1.9%	0.68
	b	Boehme	2010	BOTH	LJ AND MGIT960	LJ_PM	Peru	Peru	313	16	206	0	5	102	0	313	97.6%	100.0%	0.0%	0.67
									310		201	0	8	101	4	310	96.2%	100.0%	1.3%	0.67
	c	Boehme	2010	BOTH	LJ AND MGIT960	MGIT 960 SIRE	South Africa, Cape Town	South Africa, Cape Town	337	16	138	2	9	187	0	336	93.9%	98.9%	0.0%	0.44
									332		136	1	10	185	8	332	93.2%	99.5%	2.4%	0.44
	d	Boehme	2010	BOTH	7H11 AND MGIT 960	LJ_PM	South Africa, Durban	South Africa, Durban	264	3	43	3	2	216	0	264	95.6%	98.6%	0.0%	0.17
									261		36	3	7	215	4	261	83.7%	98.6%	1.5%	0.16
	e	Boehme	2010	BOTH	LJ_PM AND MGIT 960	MGIT 960 SIRE	India	India	224	121	182	1	6	35	0	224	96.8%	97.2%	0.0%	0.84
									222		179	0	8	35	3	222	95.7%	100.0%	1.4%	0.84
	f	Boehme	2010	BOTH	LJ, 7H11 (DURBAN), AND MGIT 960	LJ_PM, MGIT 960 SIRE	Multicountry (Azerbaijan, Peru, South Africa (2 sites), and India) laboratory validation study in	All	1356	205	710	7	30	608	2	1355	95.9%	98.9%	0.1%	0.55
									1341		675	5	57	604	23	1364	92.2%	99.2%	1.7%	0.54
2	a	Boehme	2011	BOTH	MGIT 960	MGIT 960 SIRE	Azerbaijan	Azerbaijan	536	50	203	4	26	303	37/1185	536	88.6%	98.7%	#VALUE!	0.43
	b	Boehme	2011	BOTH	MGIT 960	MGIT 960 SIRE	Peru	Peru	1005	23	171	3	6	825	18/749	1005	96.6%	99.6%	#VALUE!	0.18
	c	Boehme	2011	BOTH	MGIT 960		South Africa	South Africa	904	10	201	2	32	669	14/1195	904	86.3%	99.7%	#VALUE!	0.26
	d	Boehme	2011	BOTH	LJ and MGIT 960	Line probe	Uganda	Uganda	289	3	121	0	24	144	4/372	289	83.4%	100.0%	#VALUE!	0.50
	e	Boehme	2011	BOTH	LJ	LJ_PM	India	India	788	10	101	16	0	671	32/902	788	100.0%	97.7%	#VALUE!	0.13
	f	Boehme	2011	BOTH	Ogawa and MGIT 960	LJ_PM	The Phillipines	The Philippines	387	154	136	5	12	234	22/918	387	91.9%	97.9%	#VALUE!	0.38
	g	Boehme	2011	BOTH	LJ, Ogawa, and MGIT 960	LJ_PM, MGIT 960 SIRE, line probe assay	Multicenter (Azerbaijan, Peru, South Africa, Uganda, India, and the Philippines) implementation study in patients suspected of TB and MDR-TB, N = 6069	ALL	3909	250	933	30	100	2846	126/5321	3909	90.3%	99.0%	#VALUE!	0.26
3		Bowles	2011	BOTH	MGIT 960	7H10_AC	Laboratory-based evaluation of fresh and frozen samples (predominantly sputum) in The Netherlands, N = 89	The Netherlands	89	8	60	2	4	23	NR	89	93.8%	92.0%	#VALUE!	0.72
4		Ciftci	2011	1	LJ	Bactec 460	Laboratory-based assessment of predominatly respiratory specimens from patients suspected of TB, N = 85	Turkey	85	0	24	1	1	59	0	85	96.0%	98.3%	0.0%	0.29
5		Friedrich	2011	i	MGIT 960	MGIT 960 SIRE	Study to assess use of Xpert for patient selection process for clinical trials of anti-TB medication, N = 140, Only patients already diagnosed with TB on at least one smear entered screening	South Africa	140	3	116	1	3	6	3	129	97.5%	85.7%	2.3%	0.92
6		Hanif	2011	UNK	Solid and liquid culture	BACTEC 460	Laboratory-based assessment of respiratory specimens (predominantly sputum), N = 206, only 1 RIF resistant isolate, Kuwait	Kuwait	206	0	54	0	5	146	0	205	91.5%	100.0%	0.0%	0.29
7	a	Helb	2010	BOTH	LJ and MGIT 960		Detection MTB, clinical validation study of archived sputum samples from Vietnam, N = 107	Vietnam	107	NR	67	0	15	25	NR	107	81.7%	100.0%	#VALUE!	0.77
	b	Helb	2010	u	Culture_NR		Detection RIF rsistance, clinical validation study of archived sputum samples from retreatment TB cases in Uganda, N= 64	Uganda	64	9	63	0	1	20	NR	84	98.4%	100.0%	#VALUE!	0.76
8		Ioannidis	2011	BOTH	LJ and MGIT 960	LJ_PM and/or MGIT 960	Laboratory-based assessment of predominatly smear-negative respiratory specimens in patients suspected of TB in Greece, N = 80	Greece	80	5	29	2	3	33	2	69	90.6%	94.3%	2.9%	0.46
9		Lawn	2011	BOTH	MGIT 960	MGIT 960	Consecutive adult HIV-infected patients with no current TB diagnosis enrolling in an ART clinic	South Africa	515	4	42	2	30	322	5	401	58.3%	99.4%	1.2%	0.18
10		Malbruny	2011	BOTH	Solid culture and MGIT 960	MGIT 960	Laboratory-based assessment of patients suspected of having TB, N = 58 (predominantly bronchial aspirates 31/58), France, only 1 RIF	France	58	1	12	0	0	46	2/180	58	100.0%	100.0%	#VALUE!	0.21
11		Marlowe	2011	BOTH	Solid or liquid culture		Laboratory-based assessment of routine respiratory samples in the USA, N = 217	USA	217	NR	116	4	14	82	1	217	89.2%	95.3%	0.5%	0.60

http://systematic-review.net/

DISTILLERSR | Web-Based Systematic Review Software

Home | Features | Screenshots | The Team | Contact Us | FAQ | Pricing and Plans | Login

What is DistillerSR?

DistillerSR is an **online application designed specifically for the screening and data extraction** phases of a systematic review.

- All reviewers can work in parallel without risk of work duplication
- The system automatically handles promotion and exclusion based on your form design
- Eliminate the costly human errors associated with transcription, promotion and reference distribution. It's all handled by the system.
- Agreement reports can immediately reveal any confusion about wording
- Completion time reports allow you to make realistic timelines based on the actual time your reviewers are taking to complete forms
- Export your data to any spreadsheet or database software
- Easy protocol changes, the system can re-evaluate all user responses in seconds

That sounds great, how can we get started?

Sign up for a zero-obligation web-based demonstration and we'll be glad to show you!

GET A FREE DEMO!

Your Name*

Your Email*

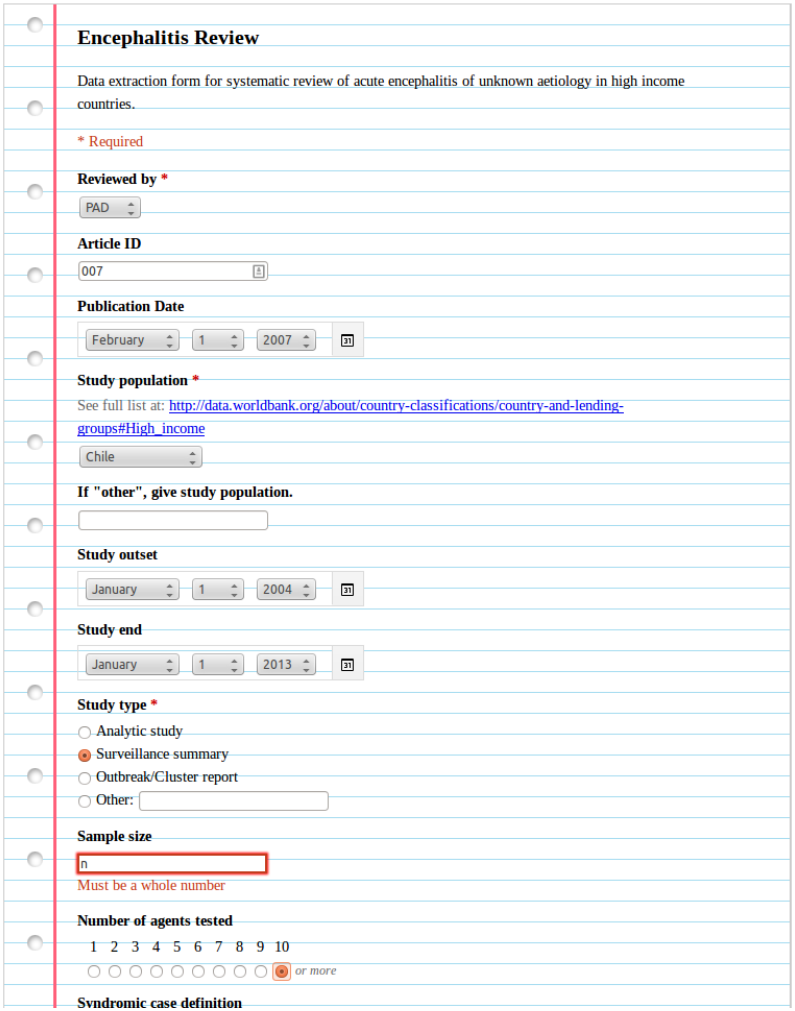
Phone Number

Please specify a preferred date and time and any other comments

Your demo should only take about 30 minutes

Sign Up!

Create a Google form



Encephalitis Review

Data extraction form for systematic review of acute encephalitis of unknown aetiology in high income countries.

* Required

Reviewed by *

PAD

Article ID

007

Publication Date

February 1 2007

Study population *

See full list at: http://data.worldbank.org/about/country-classifications/country-and-lending-groups#High_income

Chile

If "other", give study population.

Study outset

January 1 2004

Study end

January 1 2013

Study type *

- Analytic study
- Surveillance summary
- Outbreak/Cluster report
- Other:

Sample size

n

Must be a whole number

Number of agents tested

1 2 3 4 5 6 7 8 9 10 or more

Syndromic case definition

Edit this form

https://support.google.com/drive/answer/87809?hl=en (courtesy: Alex Demarsh)

Data extraction

- Process of data extraction
  - Subjective process, prone to error
  - Extraction should be preceded by training and standardization among the two reviewers
  - To make it less subjective, best done by two independent reviewers
    - If done by one person, a proportion of the forms can be cross-checked by the second reviewer
    - Disagreements between reviewers should be resolved by consensus or by a third reviewer
    - Mark the disagreements and consensus data on any one form in red ink
    - Compute inter-rater reliability using Kappa for main outcomes and study quality data

Data extraction

- Process of data extraction
  - Once data extraction is over, enter data into a database manager
    - Access, Excel, etc.
  - Enter only the consensus data
  - Compute and report inter-rater reliability
  - Missing data may be handled by contacting authors
    - Use email a lot!
    - Ask specific, pointed questions
    - Do not overwhelm the authors with too many questions!
    - Keep track of how many you contact and how many respond

For sample data extraction forms and quality assessment checklists, and author contact template, check the USB drive!

JORGE CHAM © 2008

NINJAS vs PROFESSORS

A COMPARATIVE ANALYSIS

NINJAS	PROFESSORS
Experts in methods of subterfuge	Experts in methods no longer used
Employs assortment of lethal weapons	Employs a bunch of lazy peons (you)
Can kill you without remorse	Can kill your career or worse
Always shown wearing the same outfit	Always wears the same outfit
Wears a hood	Wears a hood at graduation
Hurls Shurikens	Hurls when you present your research
People think they're pretty cool	They think they're pretty cool
Shrouded in mystery	Shrouds you in misery

WWW.PHDCOMICS.COM