Introduction to meta-analysis

Steve Simon

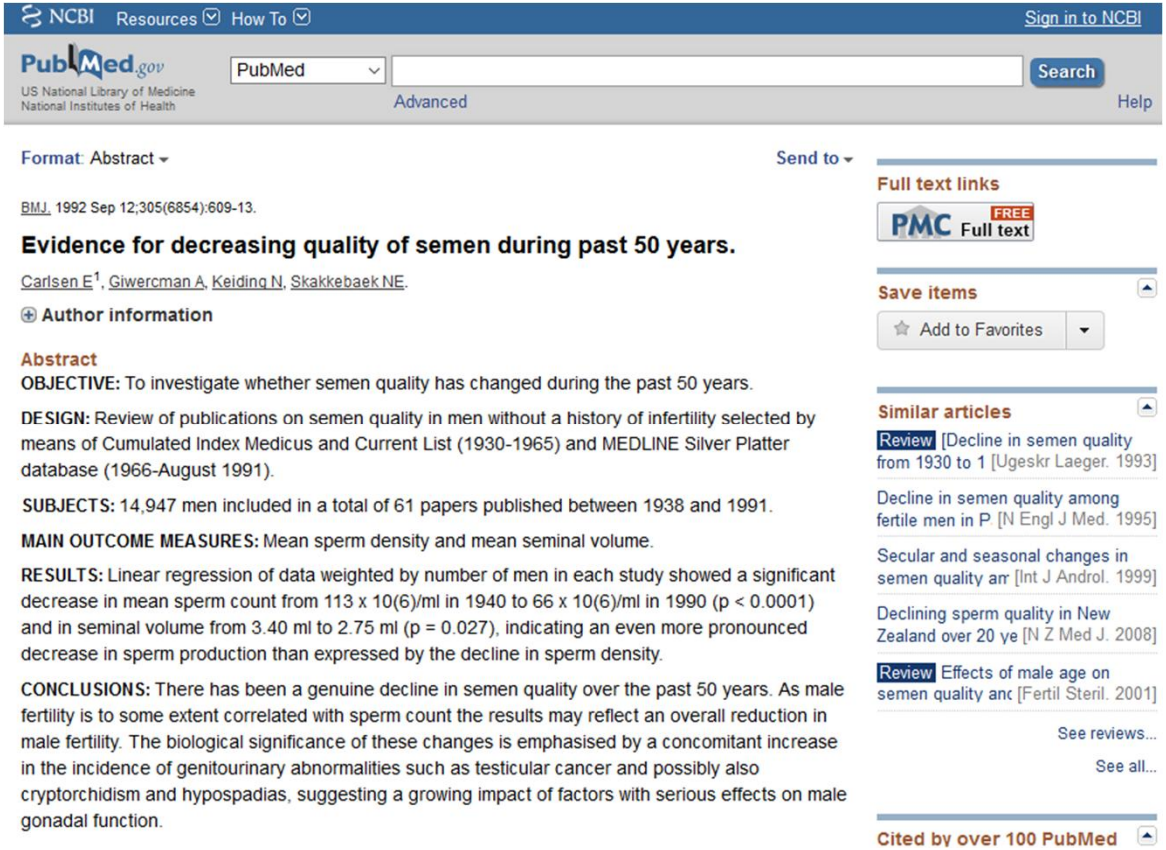
Abstract

Meta-analysis is the quantitative pooling of data from multiple studies. The three threats to the validity of a meta-analytic finding are heterogeneity, publication bias, and poor individual study quality. This talk will introduce you to the major design issues that you must address in your research protocol to insure that your meta-analysis will have credibility.

Abstract (continued)

You will also learn some of the fundamental graphical and analytic tools used in meta-analysis: the forest plot, Cochran's Q and I-squared, the funnel plot, and the L'Abbe plot. You will compare the results from a fixed effects and a random effects model and understand the choices available for summary statistics. Finally, you will see how to publish your results using the PRISMA guidelines.

A motivating example



NCBI Resources How To Sign in to NCBI

PubMed.gov
US National Library of Medicine
National Institutes of Health

PubMed Search
Advanced Help

Format: Abstract Send to

BMJ. 1992 Sep 12;305(6854):609-13.

Evidence for decreasing quality of semen during past 50 years.

Carlsen E[1], Giwercman A, Keiding N, Skakkebaek NE.

Author information

Abstract

OBJECTIVE: To investigate whether semen quality has changed during the past 50 years.

DESIGN: Review of publications on semen quality in men without a history of infertility selected by means of Cumulated Index Medicus and Current List (1930-1965) and MEDLINE Silver Platter database (1966-August 1991).

SUBJECTS: 14,947 men included in a total of 61 papers published between 1938 and 1991.

MAIN OUTCOME MEASURES: Mean sperm density and mean seminal volume.

RESULTS: Linear regression of data weighted by number of men in each study showed a significant decrease in mean sperm count from 113 x 10(6)/ml in 1940 to 66 x 10(6)/ml in 1990 ($p < 0.0001$) and in seminal volume from 3.40 ml to 2.75 ml ($p = 0.027$), indicating an even more pronounced decrease in sperm production than expressed by the decline in sperm density.

CONCLUSIONS: There has been a genuine decline in semen quality over the past 50 years. As male fertility is to some extent correlated with sperm count the results may reflect an overall reduction in male fertility. The biological significance of these changes is emphasised by a concomitant increase in the incidence of genitourinary abnormalities such as testicular cancer and possibly also cryptorchidism and hypospadias, suggesting a growing impact of factors with serious effects on male gonadal function.

Full text links
PMC FREE Full text

Save items
Add to Favorites

Similar articles
Review [Decline in semen quality from 1930 to 1 [Ugeskr Laeger. 1993]
Decline in semen quality among fertile men in P [N Engl J Med. 1995]
Secular and seasonal changes in semen quality an [Int J Androl. 1999]
Declining sperm quality in New Zealand over 20 ye [N Z Med J. 2008]
Review Effects of male age on semen quality anc [Fertil Steril. 2001]
See reviews...
See all...

Cited by over 100 PubMed

Pubmed abstract of a controversial meta-analysis

A motivating example

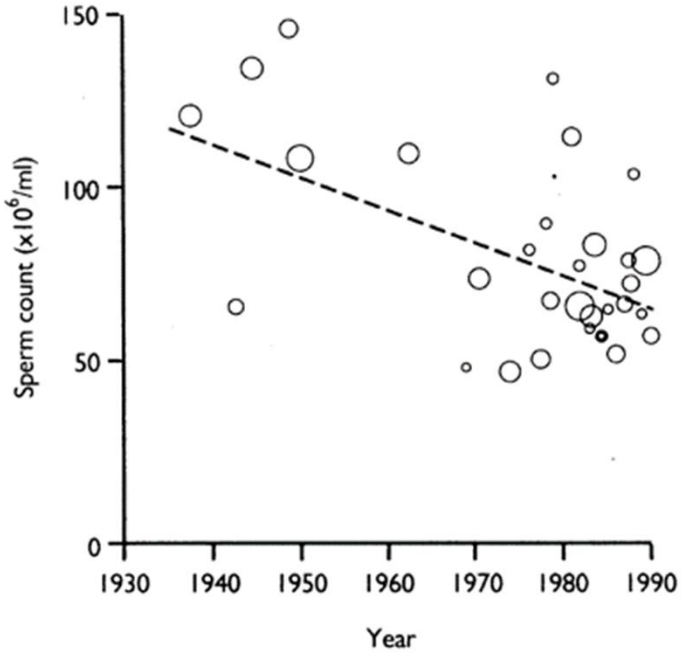


FIG 1—*Linear regression of mean sperm density reported in 61 publications (represented by circles whose area is proportional to the logarithm of the number of subjects in study) each weighted according to number of subjects, 1938-90*

graph showing decline in sperm counts over time

Alternative analysis

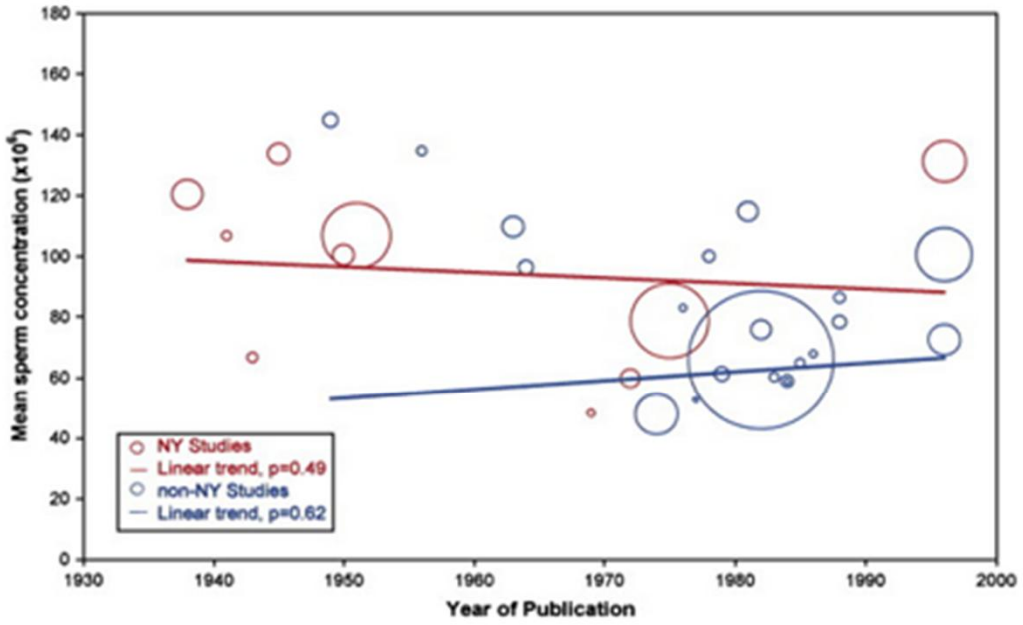


Fig. 1.

Re-analysis of data from Carlsen et al shows no decline in sperm counts (*blue regression line*) when data from New York are excluded. Bubble size corresponds to number of men in study. (*From* Saidi J, Chang D, Goluboff E, et al. Declining sperm counts in the United States? A critical review. J Urol 1999;161:460–2; with permission.)

Alternative display of semen studies

Additional sources of variation

- The source of patients
  - Sperm donor clinics,
  - Fertility work-ups
  - In vitro fertilization clinics
- Request for minimum abstinence time.
- Tobacco and marijauna use

Illustration of strengths and weaknesses of meta-analysis

- Weakness
  - Mixing North American apples with third world oranges.
- Strengths
  - Objective process.
  - Ability to re-analyze.

Apples and oranges

Here are the main threats to the validity of a meta-analysis:

- Did you mix apples and oranges? (heterogeneity)
- Did you leave some apples on the tree? (publication bias)
- Did the pile of apples amount to more than just a hill of beans? (no practical significance)
- Were all of the apples rotten? (poor study quality)

Did you mix apples and oranges?

Meta-analysis: a multi-center clinical trial where each center uses a different protocol.

How do clinical trials differ?

- How the patient population was selected.
- How the intervention was administered.
- How the controls were selected/treated.
- How the effectiveness of the intervention was measured.

You can examine heterogeneity using

- the forest plot
- L'Abbe plot
- Cochran's Q
- I-squared
- sensitivity/subgroup analysis
- meta regression

Did you leave some apples on the tree?

Publication bias is difficult to assess and difficult to control for. You should

- have a comprehensive search protocol
  - non-Medline indexed journals
  - conference presentations
  - clinical trial registries
- assess publication bias using a funnel plot.

Did the pile of apples amount to more than just a hill of beans?

Very few meta-analytic studies address practical significance

- Summary measures in meta-analysis are unitless.
- Translate your findings to a meaningful scale.

Were all of the apples rotten?

- Jadad (0-5 points total)
  - randomized? (0-2 points)
  - blinding (0-2 points)
  - information on dropouts (0-1 points)
- PEDro (0-11 points total)
  - randomization and baseline balance (0-3 points)
  - blinding (0-3 points)
  - dropouts, intention to treat (0-2 points)
  - completeness of reporting (0-3 points)

Design of a meta-analytic study

Detailed protocol

- Get help from a librarian
- Search strategy
- Inclusion/exclusion criteria
- Process for extracting numerical results

Data analysis

- Pick a summary measure
- Forest plot,
- Cochran’s Q and I-squared,
- Funnel plot
- L’Abbe plot.
- Fixed versus random effects
- Meta regression

Software

- R: see the CRAN Task View: Meta-Analysis
- SAS: metaanal macro, various SAS conference presentations
- Stata: see Meta-Analysis in Stata: An Updated Collection from the Stata Journal, Second Edition

BCG description

"Results from 13 studies examining the effectiveness of the Bacillus Calmette-Guerin (BCG) vaccine against tuberculosis."

Available at the metafor project page in html format.

BCG data, first three columns

```
##   trial                author year
## 1     1              Aronson 1948
## 2     2     Ferguson & Simes 1949
## 3     3      Rosenthal et al 1960
## 4     4    Hart & Sutherland 1977
## 5     5 Frimodt-Moller et al 1973
## 6     6      Stein & Aronson 1953
```

BCG data, last six columns

	tpos	tneg	cpos	cneg	ablat	alloc
1	4	119	11	128	44	random
2	6	300	29	274	55	random
3	3	228	11	209	42	random
4	62	13536	248	12619	52	random
5	33	5036	47	5761	13	alternate
6	180	1361	372	1079	44	alternate

LOS description

“Results from 9 studies on the length of the hospital stay of stroke patients under specialized care and under conventional/routine (non-specialist) care.”

Available at the metafor project page in html format

LOS data, first two columns

```
##   study             source
## 1     1          Edinburgh
## 2     2     Orpington-Mild
## 3     3 Orpington-Moderate
## 4     4   Orpington-Severe
## 5     5      Montreal-Home
## 6     6  Montreal-Transfer
```

LOS data, last six columns

```
##   n1i m1i sd1i n2i m2i sd2i
## 1 155  55   47 156  75   64
## 2  31  27    7  32  29    4
## 3  75  64   17  71 119   29
## 4  18  66   20  18 137   48
## 5   8  14    8  13  18   11
## 6  57  19    7  52  18    4
```

The standardized mean difference (SMD)

For a continuous outcome, the standardized mean difference is computed as

$$\hat{\theta}_i = \frac{\overline{X}_T - \overline{X}_C}{Estimated\ Standard\ Deviation}$$

or

$$\hat{\theta}_i = \frac{\overline{X}_C - \overline{X}_T}{Estimated\ Standard\ Deviation}$$

Estimating the standard deviation

Different estimated standard deviations

- Cohen's d (pooled standard deviation)
- Hedge's g (bias correction)
- Adjustments for heteroscedascity or pairing

Cohen's d

Uses a pooled standard deviation.

$$S_p = \sqrt{\frac{n_T S_T^2 + n_C S_C^2}{n_T + n_C}}$$

Hedge's g

Pooled standard deviation adjusted by a bias correction factor

$$J \approx 1 - \frac{3}{4df - 1}$$

Adjusting for heteroscedascity

Two choices for heteroscedascity within a study

$$\sqrt{\frac{S_T^2 + S_C^2}{2}}$$

or

$$S_C$$

Adjusting for pairing

For pairing use

$$S_D$$

if it is available, or

$$\sqrt{S_T^2 + S_C^2 - 2\hat{\rho} S_T S_C}$$

if it is not.

Analyzing proportions

- Commonly used summary measures
  - Odds ratio
  - Relative risk
  - Risk difference
- Odds ratios and relative risk are always analyzed and displayed on a log scale.

Titanic data

	Survived	Died
Male	142	709
Female	308	154

Mortality versus gender table

Odds ratio

	Survived	Died	odds
Male	142	709	709/142=4.993
Female	308	154	154/308=0.5

odds ratio = 9.986

Odds ratio calculation

Relative risk

	Survived	Died	P[Death]
Male	142	709	709/(709+142)=0.833
Female	308	154	154/(154+308)=0.333

Relative risk = 2.502

Relative risk calculation

Risk difference

	Survived	Died	P[Death]
Male	142	709	709/(709+142)=0.833
Female	308	154	154/(154+308)=0.333
			Risk difference = 0.500

Risk difference calculation

Standardized mean differences for los data

```
##   n1i m1i sd1i n2i m2i sd2i      yi     vi
## 1 155  55   47 156  75   64 -0.3552 0.0131
## 2  31  27    7  32  29    4 -0.3479 0.0645
## 3  75  64   17  71 119   29 -2.3176 0.0458
## 4  18  66   20  18 137   48 -1.8880 0.1606
## 5   8  14    8  13  18   11 -0.3840 0.2054
## 6  57  19    7  52  18    4  0.1721 0.0369
```

Odds ratios for the BCG data

```
##   tpos  tneg cpos  cneg      yi     vi
## 1    4   119   11   128 -0.9387 0.3571
## 2    6   300   29   274 -1.6662 0.2081
## 3    3   228   11   209 -1.3863 0.4334
## 4   62 13536  248 12619 -1.4564 0.0203
## 5   33  5036   47  5761 -0.2191 0.0520
## 6  180  1361  372  1079 -0.9581 0.0099
```

Fixed effects model

– For the SMD

$$V(\hat{\theta}_i) \approx \frac{1}{n_T} + \frac{1}{n_C}$$

– For log odds ratio

$$\approx \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}$$

– For both, weight is inverse of the variance

$$w_i = \frac{1}{V(\hat{\theta}_i)}$$

Fixed effects model (cont'd)

- Overall estimate

$$\hat{\theta} = \frac{\Sigma w_i \hat{\theta}_i}{\Sigma w_i}$$

- Variance of the overall estimate

$$V(\hat{\theta}) = \frac{1}{\Sigma w_i}$$

- Test statistic

$$\frac{\hat{\theta}^2}{V(\hat{\theta})} \sim \chi^2(1)$$

Cochrane's Q

– Test of homogeneity

$$Q = \Sigma w_i(\hat{\theta}_i - \hat{\theta})^2$$

– Distribution

$$Q \sim \chi^2(k-1)$$

– Alterantive: I-squared

$$I^2 = 100\frac{Q-(k-1)}{Q}$$

Random effects model

– Rough approximation

$$\hat{\tau}^2 = max(0, Q - (k - 1))$$

– Reweight the studies

$$W_i^* = \frac{1}{V(\hat{\theta}_i) + \hat{\tau}^2}$$

Random effects model for LOS data

```
Model Results:

estimate      se     zval    pval    ci.lb
ci.ub
 -0.5371  0.3087  -1.7401  0.0818  -1.1421
0.0679  .
```

Random effects model for BCG data

```
Model Results:

estimate      se     zval    pval    ci.lb
ci.ub
 -0.7452  0.1860  -4.0057  <.0001  -1.1098  -
0.3806  ***
```

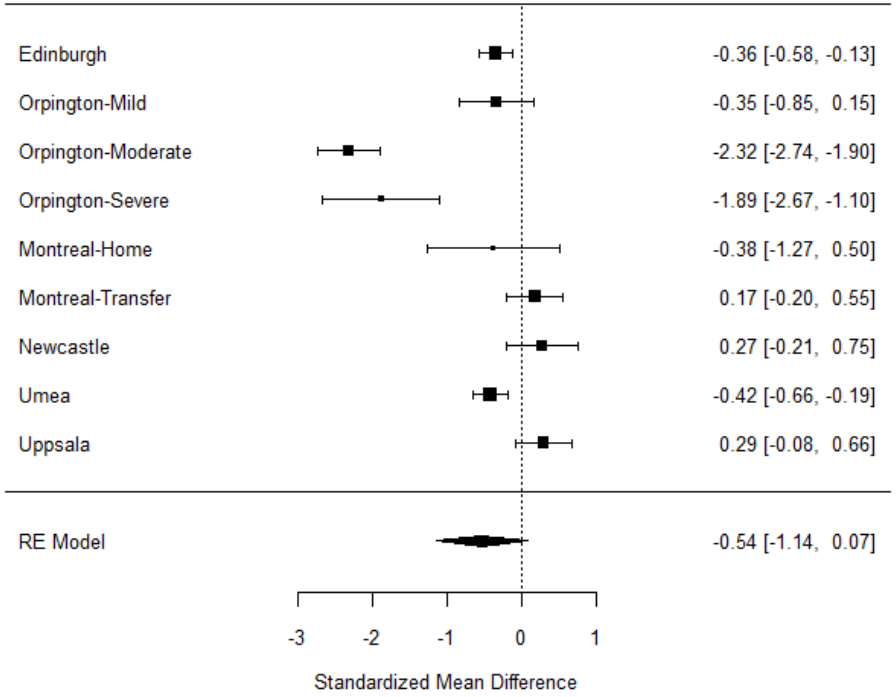
Test of homogeneity for the LOS data

```
tau^2 (estimated amount of total heterogeneity):
0.7908 (SE = 0.4281)
tau (square root of estimated tau^2 value):
0.8893
I^2 (total heterogeneity / total variability):
95.49%
H^2 (total variability / sampling variability):
22.20

Test for Heterogeneity:
Q(df = 8) = 123.7293, p-val < .0001
```

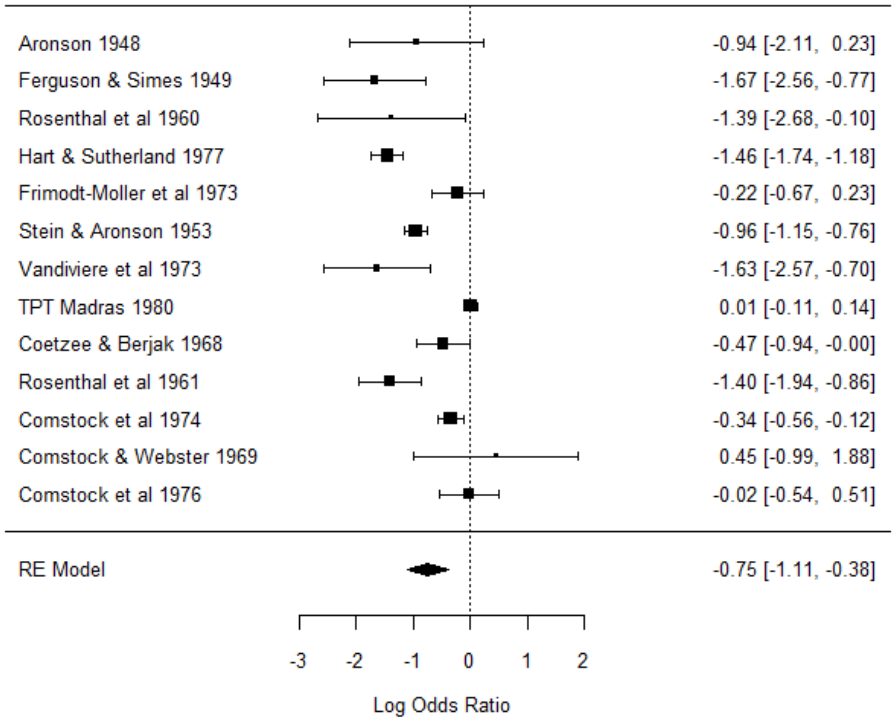
Test of homogeneity for the BCG data

```
tau^2 (estimated amount of total heterogeneity):
0.3378 (SE = 0.1784)
tau (square root of estimated tau^2 value):
0.5812
I^2 (total heterogeneity / total variability):
92.07%
H^2 (total variability / sampling variability):
12.61

Test for Heterogeneity:
Q(df = 12) = 163.1649, p-val < .0001
```

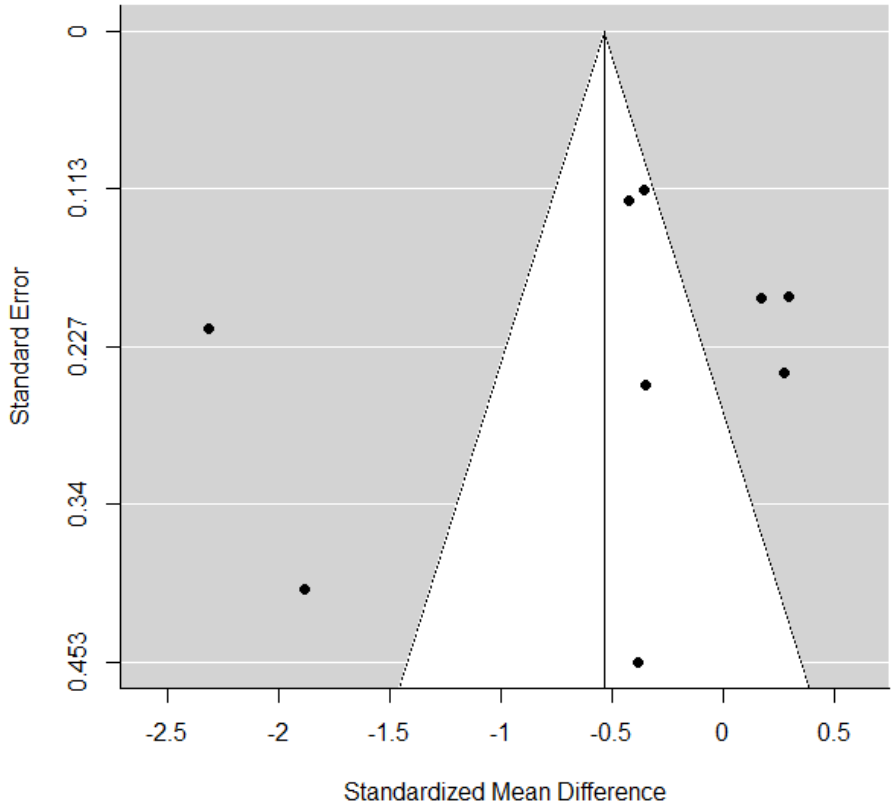
Forest plot for the LOS data



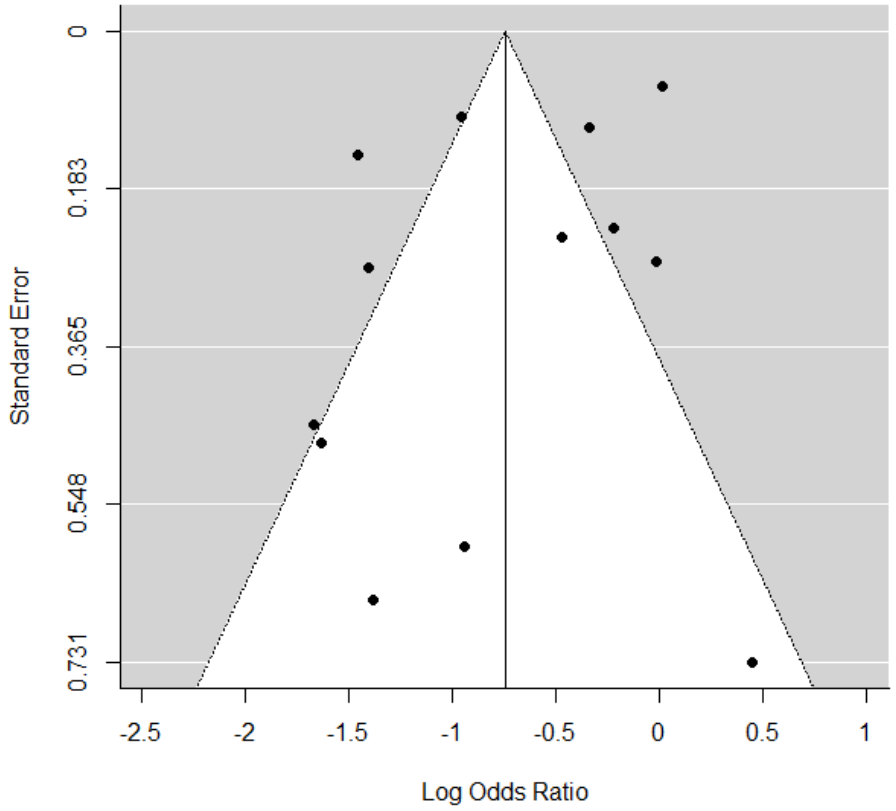
Forest plot for the BCG data



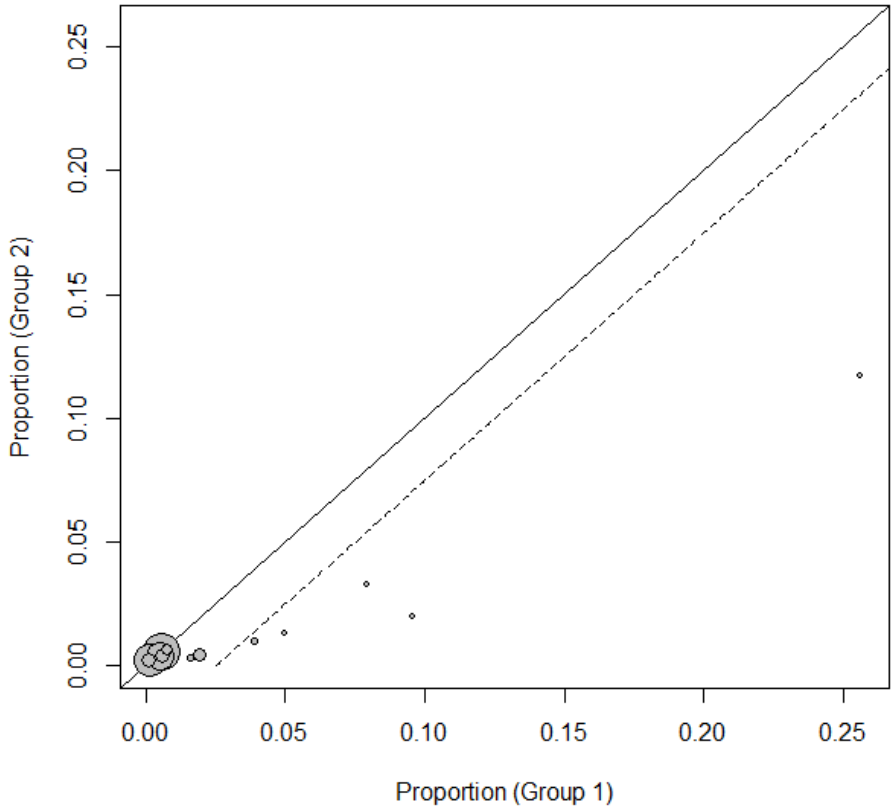
Funnel plot for the LOS data



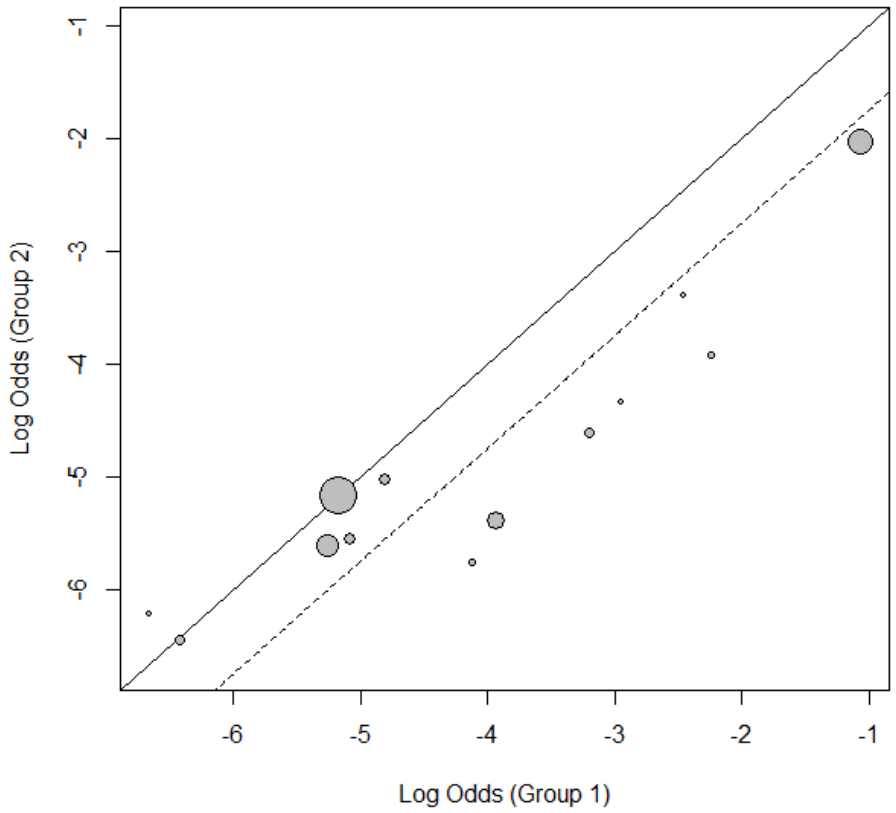
Funnel plot for the BCG data



L’Abbe plot


Proportion (Group 2)
0.00
0.05
0.10
0.15
0.20
0.25
Proportion (Group 1)
0.00
0.05
0.10
0.15
0.20
0.25

L'Abbe plot on log odds scale



Meta regression

```
Test of Moderators (coefficient(s) 2:3):
QM(df = 2) = 16.2533, p-val = 0.0003

Model Results:

         estimate       se     zval    pval
ci.lb    ci.ub
intrcpt  -10.5347  27.3739  -0.3848  0.7004  -
64.1865  43.1172
ablat     -0.0288   0.0095  -3.0311  0.0024   -
0.0475  -0.0102  **
year       0.0055   0.0138   0.3949  0.6929   -
0.0216   0.0325
```

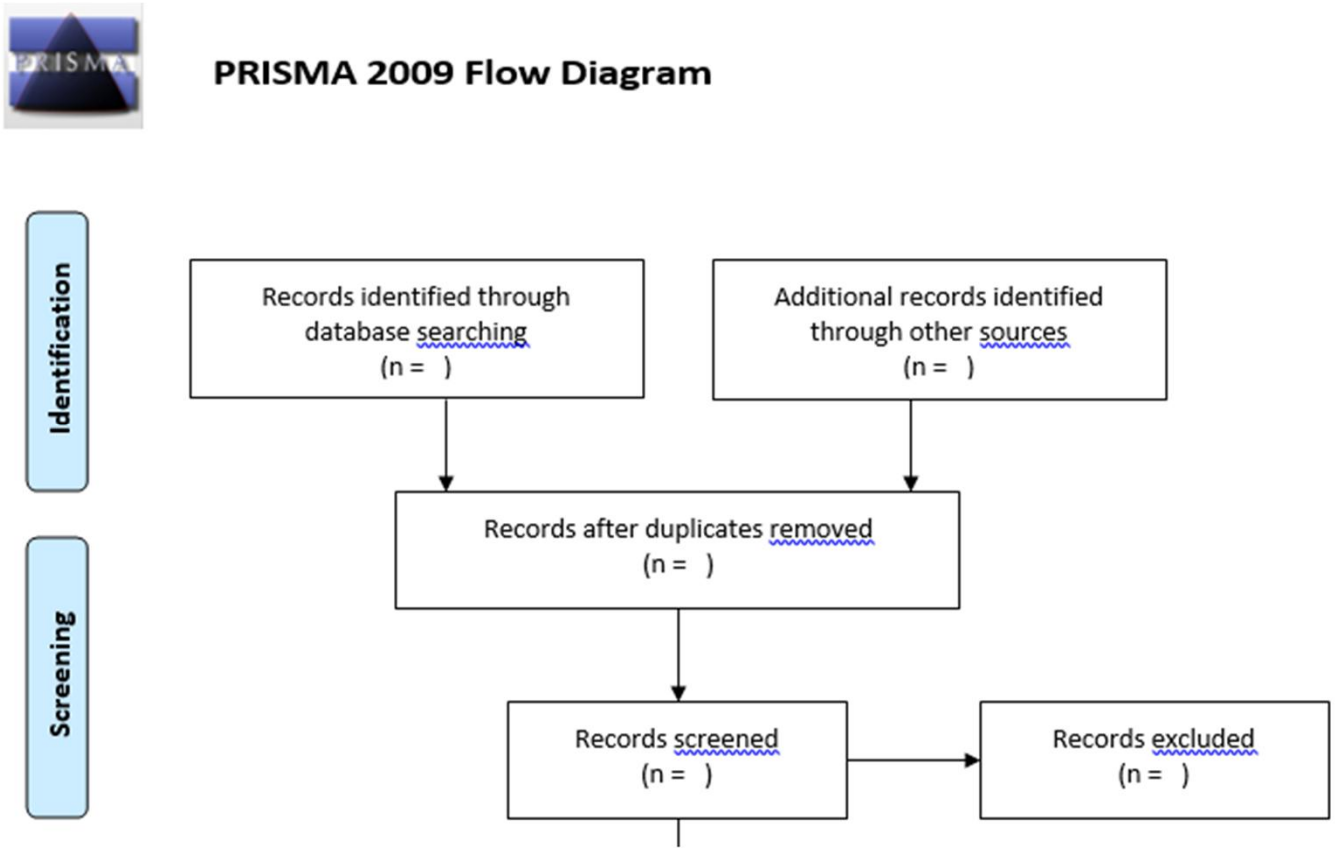
Change in heterogeneity for meta regression model

```
tau^2 (estimated amount of residual heterogeneity):     0.0913 (SE = 0.0745)
tau (square root of estimated tau^2 value): 0.3022
I^2 (residual heterogeneity / unaccounted variability): 67.29%
H^2 (unaccounted variability / sampling variability):   3.06
R^2 (amount of heterogeneity accounted for): 72.96%

Test for Residual Heterogeneity:
QE(df = 10) = 25.0121, p-val = 0.0053
```

Publication guidelines

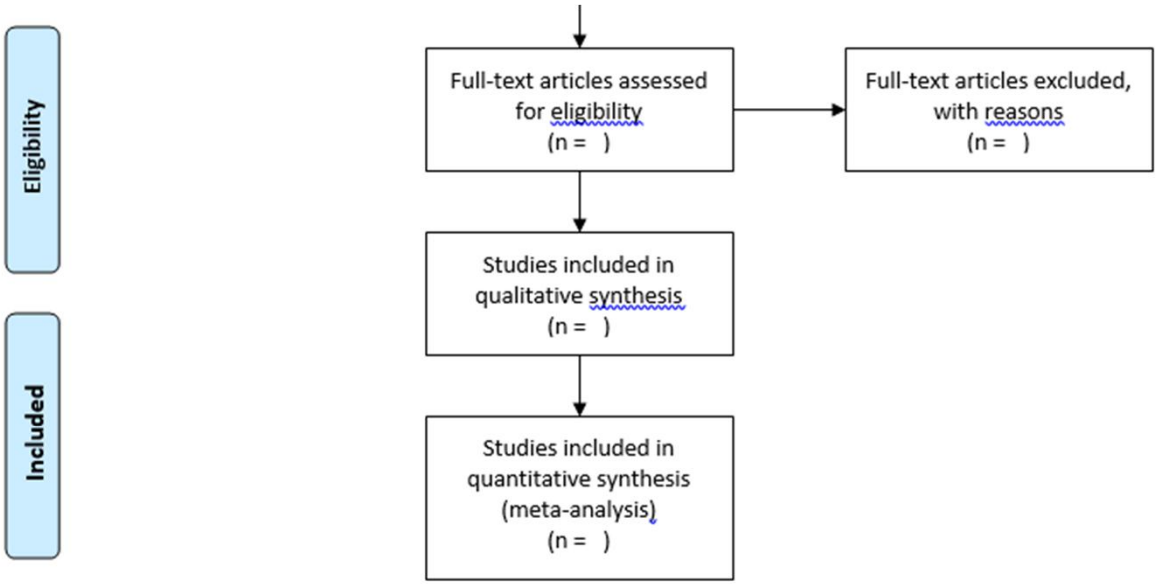
- PRISMA (2009)
  - Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- QUOROM (1996)
  - QUality Of Reporting Of Meta-analysis

PRISMA flow diagram



Top half of PRISMA flow diagram

PRISMA flow diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit www.prisma-statement.org.

Bottom half of PRISMA flow diagram

PRISMA checklist, part 1

Section/topic	#	Checklist item
TITLE		
Title	1	Identify the report as a systematic review, meta-analysis, or both.
ABSTRACT		
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.
INTRODUCTION		
Rationale	3	Describe the rationale for the review in the context of what is already known.
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).
METHODS		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.

Checklist for PRISMA

PRISMA checklist, part 2

Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.

Checklist for PRISMA

PRISMA checklist, part 3

Section/topic	#	Checklist item
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.
RESULTS		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.

Checklist for PRISMA

PRISMA checklist, part 4

Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).
DISCUSSION		
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.
FUNDING		
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Me
PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Checklist for PRISMA

Conclusion

- Design of meta-analysis
  - Include a librarian
  - Aggressive search strategy
- Statistical summary
  - Random versus fixed effects versus meta regression
  - Forest, funnel, L'Abbe plots
  - Cochrane's Q and I-squared
- Reporting
  - PRISMA flow diagram and checklist
- https://github.com/pmean/introduction-meta-analysis