Systematic reviews and meta-analyses

Jouko Miettunen, professor

Center for Life Course Epidemiology and Systems Medicine

University of Oulu

jouko.miettunen@oulu.fi

Oulu 8.4.2015

Contents

- Background

- Literature Search

- Reporting

- Criticism

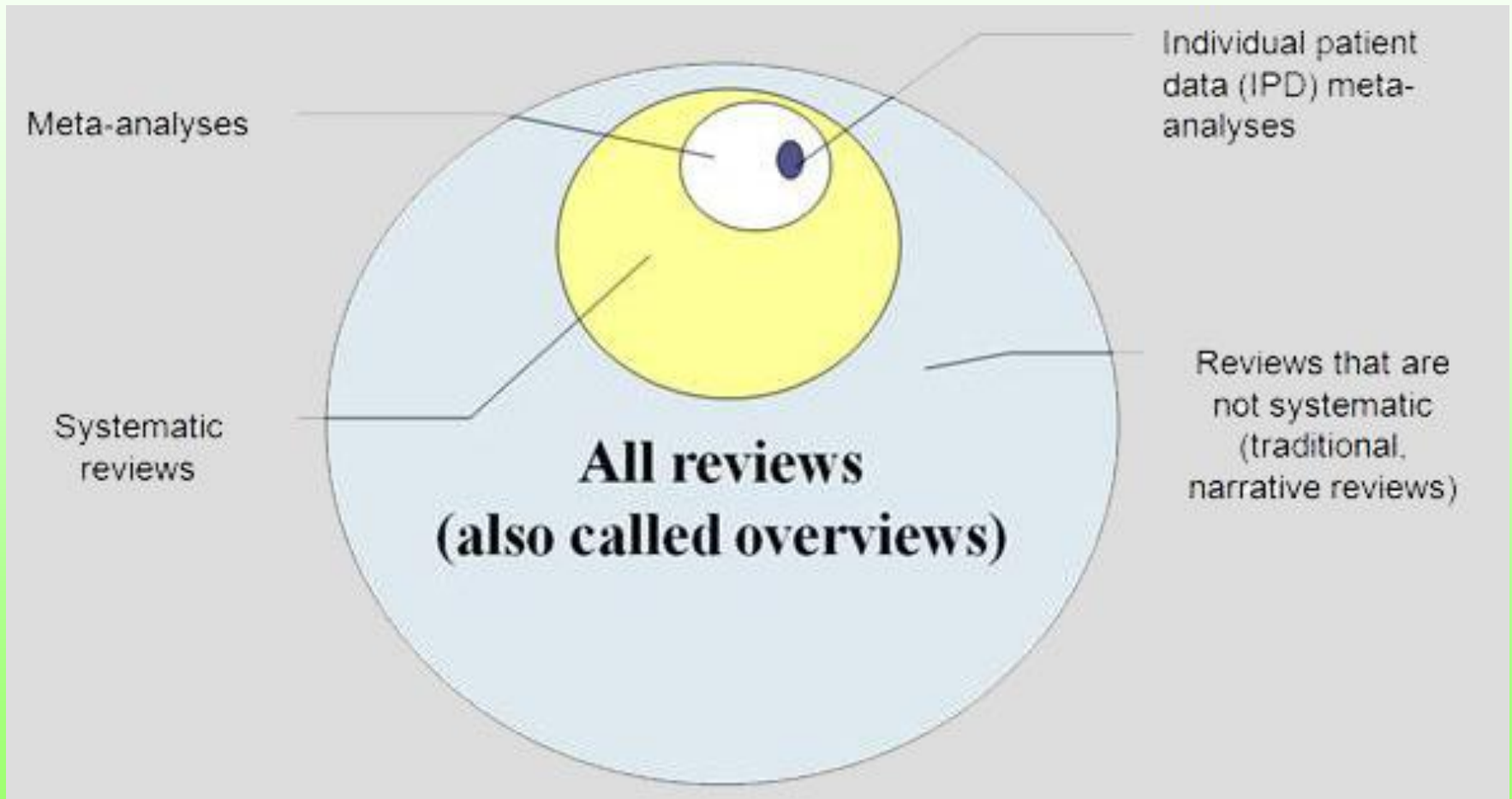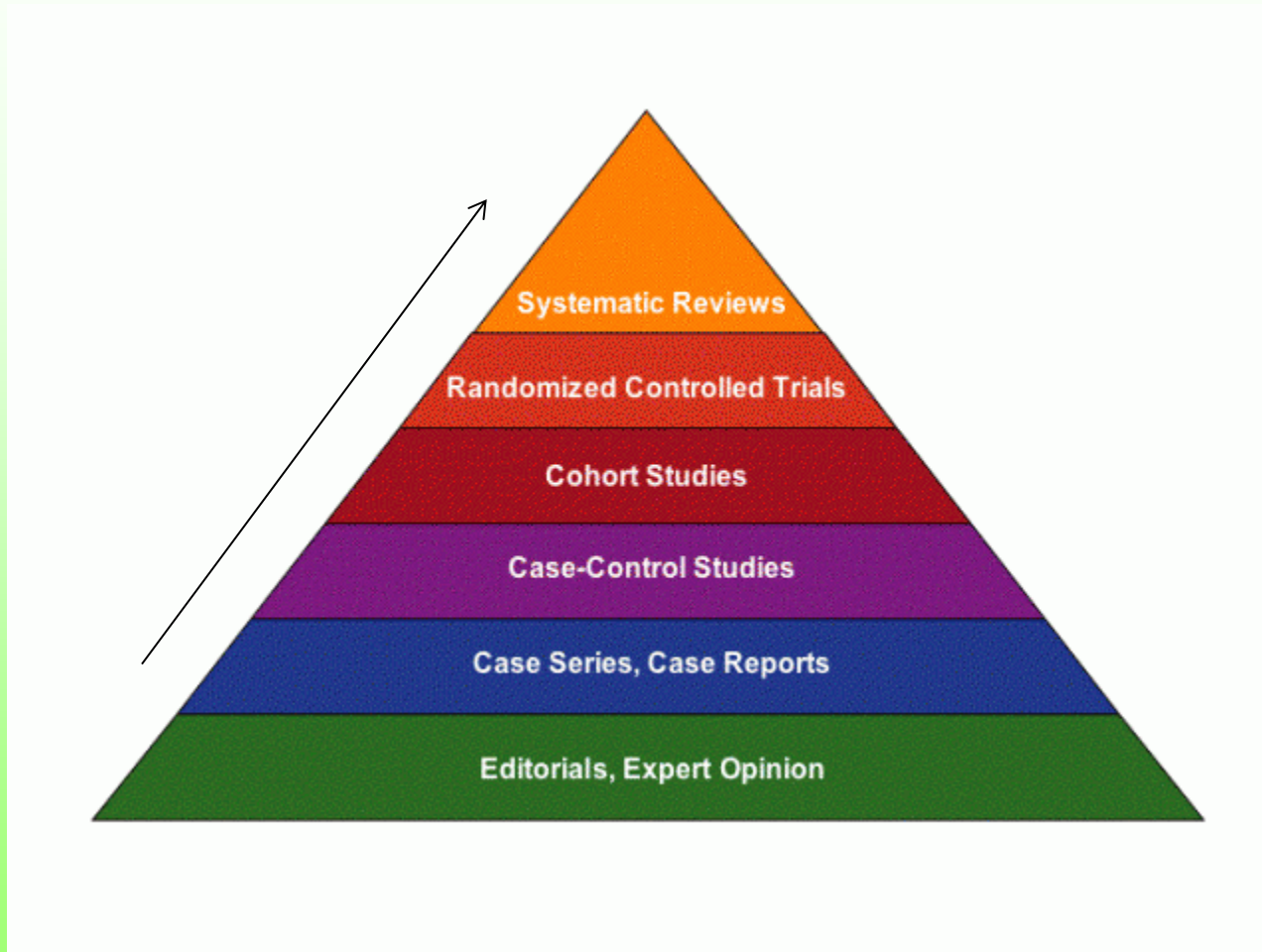- Statistical methods

- References

BACKGROUND

Systematic reviews and meta-analyses

- Review

- Systematic review

- Meta-analysis

- Pooled analysis

Systematic reviews and meta-analyses

Study designs by quality of evidence

Systematic Reviews

Randomized Controlled Trials

Cohort Studies

Case-Control Studies

Case Series, Case Reports

Editorials, Expert Opinion

Source: Evidence-Based Nursing

Systematic review

- Clearly defined research question, aim to include all studies meeting the inclusion criteria

- Methods should be replicable

- A systematic search, which is likely to found all studies that meet the criteria

- Assess the reliability of the study and potential biases

- A systematic presentation of the results

- Systematic synthesis of the characteristics and findings of the studies

- Can begin as a systematic review, one can decide later if it is going to be a meta-analysis

Meta-analyses

- Original studies are published increasingly, about 2 million medical articles per year

- Impossible to keep up-to-date, there is a need for the syntheses

- A meta-analysis combines statistically the results of previous studies

- Most often used for clinical trials (e.g. drug or therapy)

First meta-analyses

- The first statistical methods developed by Pearson (1904)

- Davis (1975) studied the antipsychotics as relapse inhibitors

- Eysenck suggested in 1952 that there is no evidence of psychotherapy functionality, this led to a heated discussion

- Smith & Glass combined in 1977 the previous 375 studied the effect of psychotherapy articles and a summary of psychotherapy that really works

Pearson K. Report on certain enteric fever inoculation statistics. British Medical Journal 1904; 3:1243-1246.
Smith & Glass. Meta-analysis of psychotherapy outcome studies. Am Psychol 1977; 32: 752-60.
Davis. Overview: Maintenance therapy in psychiatry. I. Schizophrenia. Am J Psychiatry 1975; 132:1237-45

9

Meta-analyses of clinical trials

- By far the largest part of the meta-analyses are experimental studies

- Clinical trials, interventions, and others

- The Pubmed search "meta-analysis" AND "trials" in September 2009 resulted in more than 18000, and in April 2015 about 40 000 hits

- Meta-analyses are often bases for treatment guidelines

- Cochrane meta-analyses
  - meta-analyses designed and conducted using agreed quality criteria

The Cochrane Collaboration

Working together to provide the best evidence for health care

to *The Cochrane Library*

| Home | About us | Cochrane Reviews | News & Events | Training | Multimedia | Contact | Community |

🔎 Search site　go

About us

About us

- Our principles
- ▶ History
- Evidence-based health care
- ▶ Webliography of evidence-based health care resources
- ▶ Relations with the World Health Organization
- Newcomers' guide
- Frequently Asked Questions (FAQ)
- Funding & support
- How to donate
- Get involved
- ▶ Our policies
- ▶ Reports
- ▶ Awards, scholarships & funding initiatives
- The Campbell Collaboration

About us

❝**Our vision** is that healthcare decision-making throughout the world will be informed by high-quality, timely research evidence.❞

The Cochrane Collaboration, established in 1993, is an international network of more than 28,000 dedicated people from over 100 countries. We work together to help health care providers, policy-makers, patients, their advocates and carers, make well-informed decisions about health care, based on the best available research evidence, by preparing, updating and promoting the accessibility of Cochrane Reviews – over 5,000 so far, published online in *The Cochrane Library*.

Annual Report

Annual Report & Financial Statements 2010/11

Visit the Annual Report Online Portal

Thanks to our funders

Australia, Children's Hospital at Westmead

⏸ ⏭

see all funders

11

Meta-analyses of clinical trials

Examples

- Tamoxifen in the treatment of breast cancer

- Aspirin in the treatment of cardiovascular diseases

- Antibiotics in the treatment of urinary tract infections in children

Bartolucci. Yonsei Med J 2007;48:157-63.

Meta-analyses of observational studies

- Meta-analyses of observational studies can examine risk factors (eg. passive smoking & lung cancer), prevalence of a disease, etc.

- Saha et al. (PLoS Medicine 2005; 2:e141) estimated the schizophrenia life-time prevalence to be 0.4-0.7%

- Polanczyk et al. (Am J Psychiatry 2007; 164:942-8) estimated the ADHD prevalence to be 5.6 %

FIGURE 2. ADHD/HD Pooled Prevalence According to Demographic Characteristics and Geographic Location

Polanczyk ym. Am J Psychiatry 2007; 164: 942-8
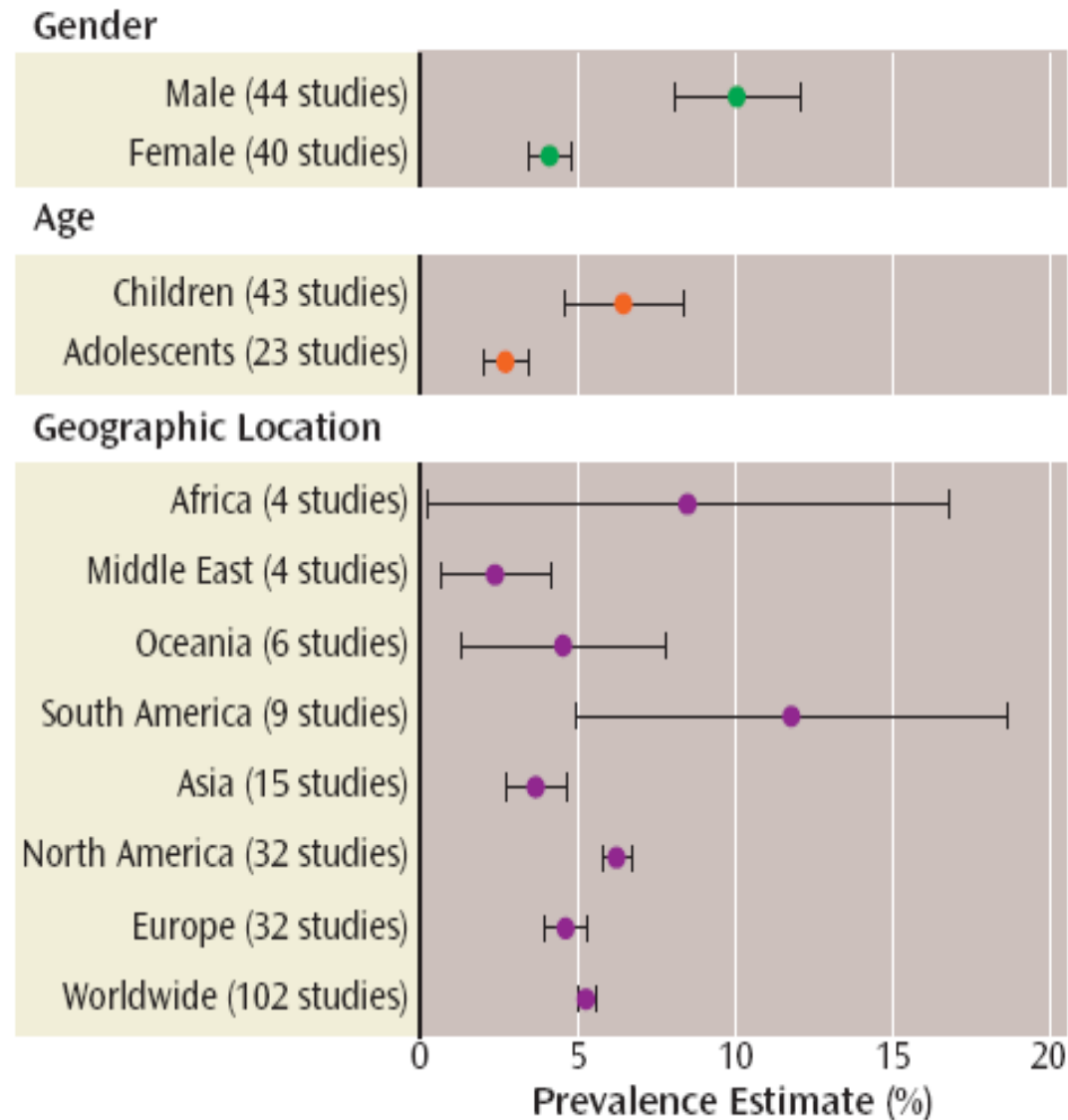
Meta-analyses of observational studies

- Inter-correlations between temperament traits
- Cultural differences in temperament traits
- Temperament in psychiatric illnesses when compared to controls
- The prevalence of alcoholism in schizophrenia
- Gender differences in schizotypal traits
- Recovery rate in schizophrenia
- Association between family history to outcome of schizophrenia

LITERATURE SEARCH

Research question?

- A clear and limited question or questions

- Unclear association or effect?

- The magnitude of effect is unclear?

- Original studies were not powered enough?

- Heterogeneity of findings and factors behind it?

Research question?

PICO(TS)

- **P**opulation (participants)
- **I**ntervention
- **C**omparator (controls/placebo)
- **O**utcome
- **T**ime (duration)
- **S**tudy design (experimental / observational, N)

- Inclusion/exclusion criteria
  - ◆ Sample size, diagnostics, trial...
  - ◆ rather wide than narrow criteria
  - ◆ easier to subsequently limit the study

Systematic search

- Multiple databases
  - removal of duplicates
- Unpublished studies
  - clinical trial databases
  - conference abstracts
  - contact authors
- Manual search
  - reference lists
  - books

Systematic search

- Selection of keywords

  - synonyms, access to logical operators

  - e.g. ((schizophrenia OR psychosis) AND (cognition OR brain)) NOT (intervention OR trial)

- Depending on the resources?

- Time criteria? Language? Full text?

- Search limited to title or abstract?

- MeSH terms?

- Search Criteria, databases, and the search date should be indicated

  - update if needed!

Appendix DS1   An example of complete search strategy, performed in Scopus.

1. ALL("duration of untreated psychosis")
2. ALL("delay in treatment")
3. ALL("treatment delay")
4. ALL("initiation of treatment")
5. ALL("duration of untreated illness")
6. #1 OR #2 OR #3 OR #4 OR #5
7. ALL(psychosis)
8. ALL("psychotic disorders")
9. ALL(schizophrenia)
10. ALL(schizoaffective)
11. ALL(schizophreniform)
12. #7 OR #8 OR #9 OR #10 OR #11
13. #6 AND #12
14. #13 AND LIMIT-TO(DOCTYPE, "ar")
15. #13 AND LIMIT-TO(DOCTYPE, "re")
16. #14 OR #15

21

Selection of databases

Table I. Key medical bibliographic databases for literature search.

CINAHL (The Cumulative Index to Nursing & Allied Health Literature)	Elsevier ScienceDirect	EMBASE (Excerpta Medica)	LILACS (Latin American and Caribbean Literature on the Health Sciences)	Medline (Ovid) (Medical Literature Analysis and Retrieval System Online)
- Bibliographic nursing and healthcare database of Ovid Technologies. - About 1.5 million references to articles, congress publications and academic dissertations since 1982. - About half of the references are found in the PubMed database (20).	- A database maintained by Elsevier B.V. containing bibliographic data and full texts. - About 6.75 million articles up to 1995 and 2.75 million articles from 1994 onwards. - Covers 25% of full texts and bibliographic data in science, technology and medicine in the world (21).	- A bibliographic biomedical and pharmacological database produced by Elsevier B.V. - Over 11 million records from 5,000 magazines from 1974 onwards. - More than 500,000 references and abstracts are added to the database each year (22).	- Open-access health science database of BIREME Systems in Spanish, Portuguese and English. - About 150,000 records, such as books, congress and conference publications, and articles from 670 well-known medical journals (23).	- Bibliographic database published by Ovid Technologies. - About 13 million references on medicine and related fields from 4,800 magazines since 1966. - An increasing number of references contain a link to freely available full text (24).

Löhönen et al. Int J Circumpolar Health 2009; 68: 394-404

PsycINFO	PSYNDEX	PubMed	Scopus	Web of Science
- A bibliographic psychological database provided by EBSCO Publishing.	- A bibliographic psychological database from the German-speaking countries.	- A free service of the U.S. National Library on Medicine through which also Medline is available.	- Bibliographic database of Elsevier B.V.	- Bibliographic database of Thomson Reuters.
- 2.3 million references and abstracts from year 1887.	- All areas of psychology and related behavioural and social sciences from 1977, audiovisual media from 1932, and tests from 1945 (26).	- About 16 million references from the 1950s onwards.	- About 27 million abstracts, 230 million references, 200 million scientific www-pages, over 12,850 journals, 535 of which are OA journals.	- Databases accessible from 1986 on: Science Citation Index Expanded, Social Science Citation Index, Arts & Humanities Citation Index.
- References from sources such as articles, books and academic dissertations in all fields related to psychology (25).		- Includes e.g. new references that are not yet indexed in Medline.	- Covers the Medline (Ovid) database, including full text links when applicable.	- 850,000 references including links to full texts when applicable.
		- Links to full text if the organization subscribes to the magazine in question (27)].	- Possibility to examine citedness (28).	- Possibility to examine citedness (29).

Coverage of the bibliographic databases in mental health research

JOHANNA LÖHÖNEN, MATTI ISOHANNI, PENTTI NIEMINEN, JOUKO MIETTUNEN

I. Schizotypal personality N=415
PubMed 65% (271)
5.1% (21)
PsycINFO 88% (367)
11.3% (47)
2.7% (11)
46.3% (192)
Web of Science 64% (265)
19.8% (82)
11.0% (46)
3.9% (16)
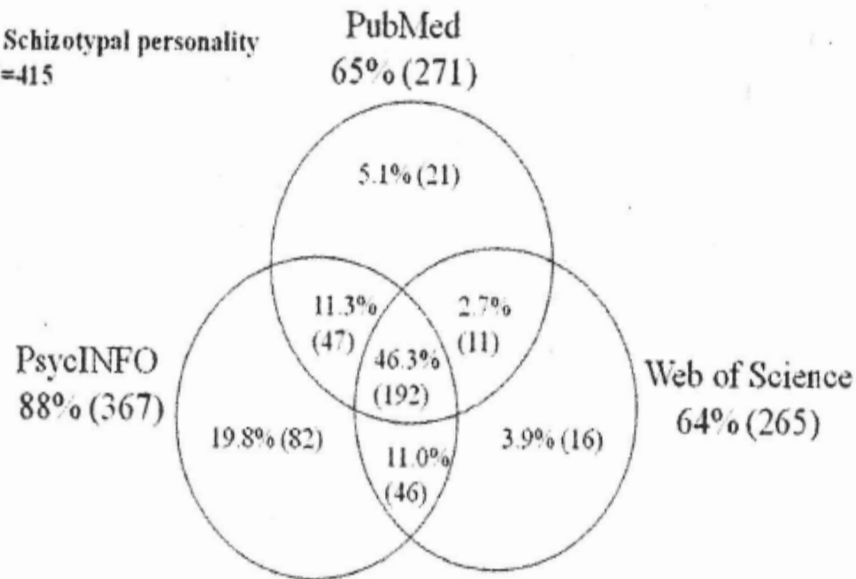
Combinations of databases
PubMed + Web of Science 80% (333)
PubMed + PsycINFO 96% (399)
PsycINFO + Web of Science 95% (394)

II. MRI N=573
PubMed 72% (410)
25.0% (143)
PsycINFO 37% (211)
3.3% (19)
18.0% (103)
25.3% (145)
Web of Science 65% (375)
6.3% (36)
1.9% (11)
20.2% (116)
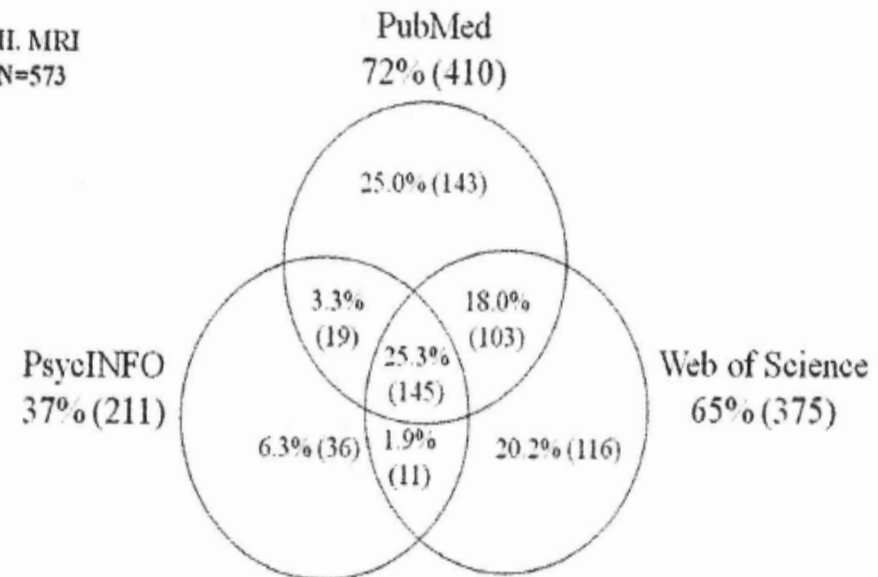
Combinations of databases
PubMed + Web of Science 94% (537)
PubMed + PsycINFO 80% (457)
PsycINFO + Web of Science 75% (430)

Evaluation of articles

- Evaluation of articles
  - ◆ title
  - ◆ abstract
  - ◆ article text
- Two reviewers evaluates all articles
  - ◆ relating to inclusion criteria
  - ◆ relating to data collection (results)
  - ◆ in difficult topics reliability of evaluators can be assessed
- Resources?

REPORTING

- Good reporting helps to assess the research

  ◆ transparency, replicability

- Clinical trials are of higher quality (stricter rules)

  ◆ Cochrane Risk of Bias -tool (Higgins & Altman, www.cochrane-handbook.org)

- Why the meta-analysis is done?

  ◆ background?

  ◆ has it been done before?

  ◆ differences to the previous meta-analysis?

Reporting guidelines

- Moher et al. 2009, Liberati et al. 2009 (PRISMA)

  - ◆ For clinical trials, may need to be edited for observational studies

- Stroup et al. 2000 (MOOSE)

  - ◆ Observational studies

- Shea et al. 2007 (AMSTAR)

  - ◆ Tool for evaluating the quality of the studies

PRISMA checklist

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).

Moher et al. 2009, Liberati et al. 2009

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.
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.
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.

Moher et al. 2009, Liberati et al. 2009

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 and (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.
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., health care 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.

Moher et al. 2009, Liberati et al. 2009

Different phases

- Search
- Evaluation
- Sources of biases
- Data
- Qualitative synthesis
- Meta-analysis

Search / Evaluation

- Number of studies provided by the search
- Number of articles after evaluation of studies based on the inclusion criteria
- Partly overlapping data sets, etc.
  - can check the name of the project, the authors, the country, the sample size, variables, etc.
- Excluded studies can be reported (table, the cause for exclusion)
- Flow chart!

Identification

# of records identified through database searching

# of additional records identified through other sources

Screening

# of records after duplicates removed

# of records screened

# of records excluded

Eligibility

# of full-text articles assessed for eligibility

# of full-text articles excluded, with reasons

Included

# of studies included in qualitative synthesis

# of studies included in quantitative synthesis (meta-analysis)
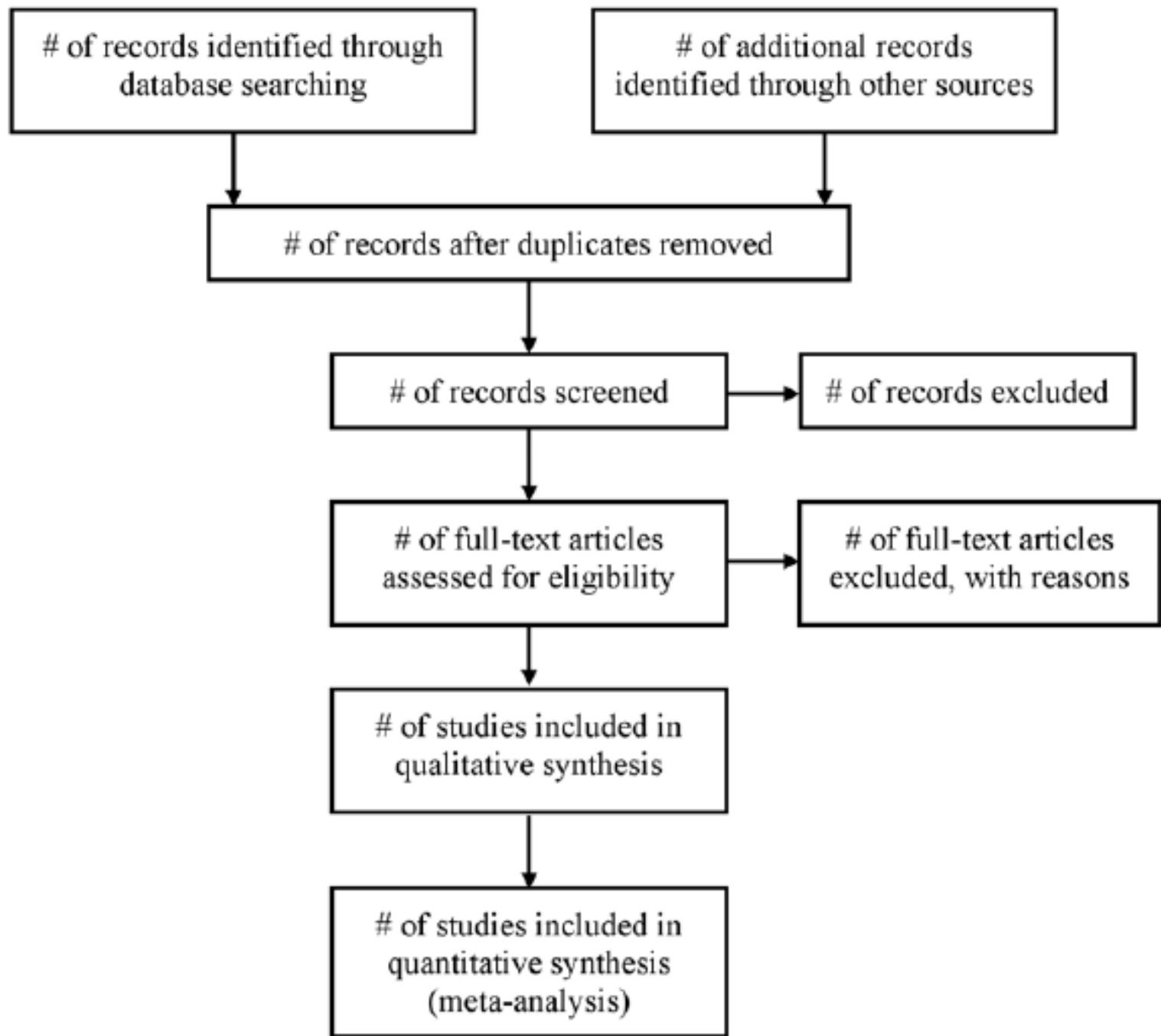
Online supplement Appendix 1. Quality Reporting Scale

Items	Quality score
Case Ascertainment (only one and highest score possible)	
Hospital inpatient OR outpatients OR case registers	2
Unspecified	0
Diagnosis (only one and highest score possible)	
Any diagnostic system reported (e.g., CATEGO, DSM, Feighner's, Schneider's, RDC, ICD, certain local guidelines (e.g. Scandinavian concept))	3
Own system described	1
Unspecified	0
Other (select from zero to six)	
Number of total sample and recovered cases unambiguously described	1
Drop out rate of schizophrenic psychoses described and rate at most 30%	1
Sample and methods clearly described (country, data collection period, follow-up period, age of the sample, diagnostics)	1
Recovery percentage presented separately for cases with schizophrenic psychoses or schizophrenia	1
Sample including outpatients OR sample including inpatients from multiple institutions	1
Results for reliability reported for diagnostics or outcomes assessments	1

Jääskeläinen et al. 2013

35

Collected variables

- form for the data collection

Online supplement Appendix 2. List of all variables for which data were sought.

Project name
Country
Data collection years
Mean follow-up time (since onset, since discharge)
Follow-up method (register, interview, other)
Diagnostic system
Patient types (age, diagnoses, outpatients vs. hospital patients, one center vs. multicenter, first episode vs. general intake)
Information on drop-outs (%, attrition analysis)
Sample size and male-female –ratio
Recovery criteria
Number and percentage of recovered subjects
Outcomes (social, clinical, combined, only course of illness)
Instruments for measuring outcomes
Analyzed predictors of outcomes
Strengths and limitations of the study
Comments

Jääskeläinen et al. 2013

Literature table

- all studies and essential data included

Supporting Table 1. Characteristics of included studies.

NAME references	Scan interval (mean)	Sample size(M/F)	Diagnostic-system/ diagnosis/ duration	MRI scanner / slice thickness	Brain areas	Used predictors	Comments, covariates
Iowa Longitudinal Study, IA, USA							
Westmoreland Corson et al., 1999	2 years	23 (23/0) mostly atypical n=13 mostly typical n=10	DSM-IV / sch (83%) / FE	1.5 T GE Signa / 1.5 mm	basal ganglia (caudate, putamen, globus pallidus)	atypical and typical cumulative dose	Subsample of the Iowa Longitudinal Study. No covariates.
Ho et al., 2003	3 years	73 (53/20)	DSM-IV / sch / FE (2 y)	1.5 T GE Signa / 1.5. mm	whole brain, lateral ventricles, sulcal CSF, total and CSF of frontal, temporal and parietal lobes, cerebellum	cumulative dose	In sample 70/73 from Iowa longitudinal study. Partly overlapping comparisons with studieswith larger samples and longer follow-up (Ho et al. 2007, Ho et al. 2011, Andreasen et al. 2013). Covariates: age, sex, height, length of inter-scan interval.

Huhtaniska et al. Manuscript.

Sources of bias

- Weaknesses / Limitations
- Original studies
  - ◆ quality
- Combining studies and review
  - ◆ search
  - ◆ publication bias (funnel plot)
  - ◆ sensitivity and subgroup analyses
- Strength of the finding can be estimated (GRADE, Guyatt 2008)
  - ◆ "meta-review"

Guyatt et al. GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336: 924–926.

Matheson SL, Shepherd AM, Carr VJ. How much do we know about schizophrenia and how well do we know it? Evidence from the Schizophrenia Library. Psychol Med 2014; 44:3387-405.

Background. True findings about schizophrenia remain elusive; many findings are not replicated and conflicting results are common. Well-conducted systematic reviews have the ability to make robust, generalizable conclusions, with good meta-analyses potentially providing the closest estimate of the true effect size. In this paper, we undertake a systematic approach to synthesising the available evidence from well-conducted systematic reviews on schizophrenia.

Method. Reviews were identified by searching Medline, EMBASE, CINAHL, Current Contents and PsycINFO. The decision to include or exclude reviews, data extraction and quality assessments were conducted in duplicate. Evidence was graded as high quality if reviews contained large samples and robust results; and as moderate quality if reviews contained imprecision, inconsistency, smaller samples or study designs that may be prone to bias.

Results. High- and moderate-quality evidence shows that numerous psychosocial and biomedical treatments are effective. Patients have relatively poor cognitive functioning, and subtle, but diverse, structural brain alterations, altered electrophysiological functioning and sleep patterns, minor physical anomalies, neurological soft signs, and sensory alterations. There are markers of infection, inflammation or altered immunological parameters; and there is increased mortality from a range of causes. Risk for schizophrenia is increased with cannabis use, pregnancy and birth complications, prenatal exposure to *Toxoplasma gondii*, childhood central nervous system viral infections, childhood adversities, urbanicity and immigration (first and second generation), particularly in certain ethnic groups. Developmental motor delays and lower intelligence quotient in childhood and adolescence are apparent.

Conclusions. We conclude that while our knowledge of schizophrenia is very substantial, our understanding of it remains limited.

Sources of bias

- Publication bias is estimated with a funnel plot

- It is assumed that the most accurate (and the largest) studies give average results, the smaller studies should be on both sides of the average

- "Trim and fill", "Fail -safe" N

Rosenberg. Evolution 2005;59: 464-8

Funnel Plot

Aspirin effect on No. of Tumor Bearing
Risque relatif, modèle fixe (IC 95%)

Corpet & Pierre. Eur J Cancer 2005 (http://corpet.free.fr/MAaspirin.html)
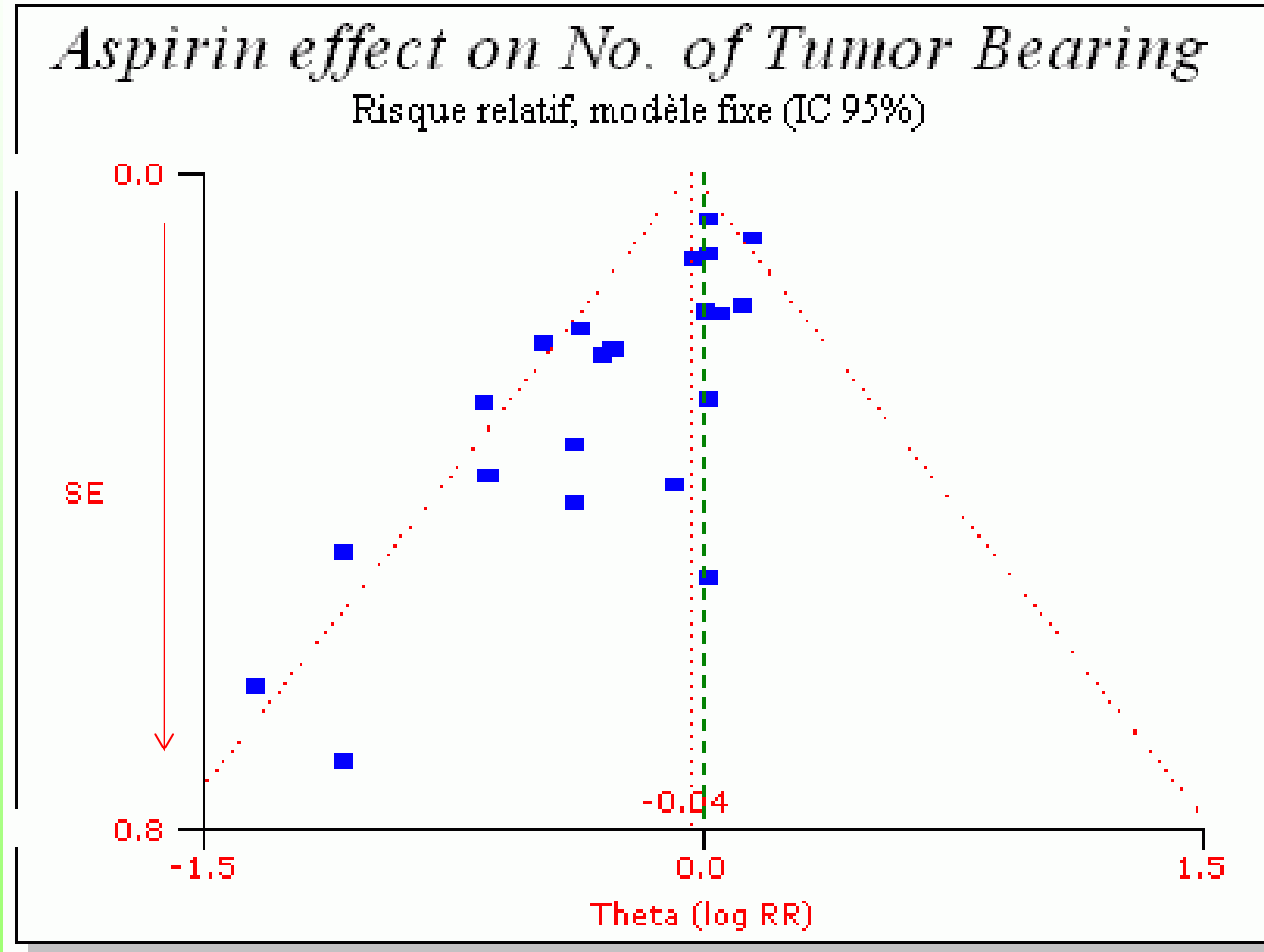
Trim and Fill

- can be used to correct for publication bias

Aspirin effect on No. of Tumor Bearing

Risque relatif, modèle fixe (IC 95%)
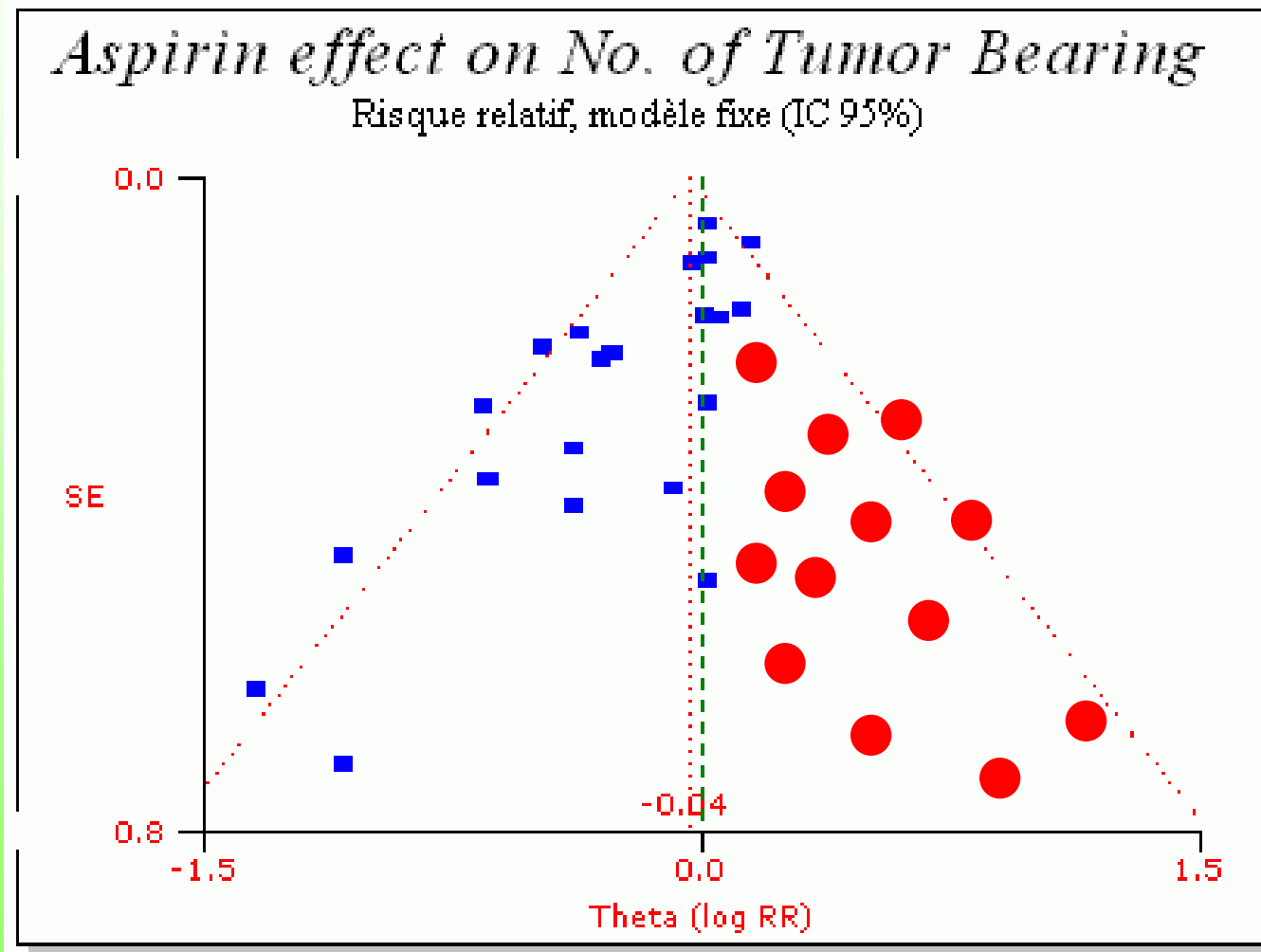
Results

- Pooled effect (95% confidence interval)

- Forest plot

- Clinical trials

  - ◆ relative vs. absolute risk reduction (Leucht et al. 2009)

Aspirin effect on No. of Tumor Bearing Rats

Trial	T. étudié n/N	T. controle n/N	Graphic	RR [95%CI]
Barnes99	15 / 32	15 / 32		1.00 [0.59; 1.68]
Miliaras04 30mg/kg/d	7 / 14	15 / 15		0.51 [0.30; 0.86]
Molck02 300ppm+Vit.D	6 / 16	9 / 16		0.67 [0.31; 1.43]
Molck02 300ppm	6 / 16	9 / 16		0.67 [0.31; 1.43]
Li & Schut99 1800ppm	16 / 30	22 / 31		0.75 [0.50; 1.12]
Li & Schut99 600ppm wk11-20 / 27		22 / 31		1.04 [0.76; 1.43]
Li & Schut99 200ppm wk11-22 / 28		22 / 31		1.11 [0.82; 1.49]
Li & Schut99 1800ppm	15 / 29	22 / 31		0.73 [0.48; 1.11]
Li & Schut99 200ppm	24 / 34	22 / 31		0.99 [0.73; 1.36]
Miliaras04 10mg/kg/d	15 / 15	15 / 15		1.00 [0.91; 1.10]
Barnes99	10 / 32	15 / 32		0.67 [0.35; 1.25]
Craven92 50	5 / 12	5 / 12		1.00 [0.39; 2.58]
Barnes99	5 / 32	15 / 32		0.33 [0.14; 0.81]
Pence95 wk2-36	35 / 43	40 / 47		0.96 [0.79; 1.15]
Pence95 wk36-51	31 / 32	40 / 47		1.14 [0.99; 1.30]
Davis94 60mg	2 / 8	8 / 8		0.26 [0.08; 0.86]
Davis94 30mg	4 / 8	8 / 8		0.52 [0.26; 1.04]
Davis94 5mg	8 / 8	8 / 8		1.00 [0.84; 1.19]
Reddy93 200ppm	19 / 36	28 / 36		0.68 [0.48; 0.97]
Reddy93 400ppm	17 / 36	28 / 36		0.61 [0.41; 0.89]
Craven92 10	2 / 12	6 / 12		0.33 [0.08; 1.33]
Li & Schut99 400ppm	9 / 27	10 / 27		0.90 [0.44; 1.86]
Global	**p ass=0.008**			0.86 [0.77; 0.96]

Het. between the 22 trials p=-9.0000 , I^2=0%

0.2   1   2.0

44

Conclusions

- Not too optimistic!

- Critical

- Between and within study biases

- Discuss possible effect of excluded studies, i.e. effect of inclusion and exclusion criteria

Criticism

Heterogeneity

- Initial studies differ significantly from each other (heterogeneous), so the pooling is problematic, e.g.
  - Differences in the assessment methods
  - Sample selection methods
  - Different adjustments in analyses or unadjusted analyses
    - ☞ covariates can be presented in the literature table
- Metaregression, median effects, ...
- The subgroup results

- The methods are difficult to understand
- The selections made by those performing meta-analysis
  - ◆ inclusions, exclusions, risk/outcome, follow-up time, etc.
- Studies do not report all the information you wished to use
- Uncertainty of the small risk estimates received?

- Although methods and results are similar… → interpretations done by the researchers may vary
- Cochrane reviews: amisulpride (Mota et al. 2002) vs. olanzapine (Duggan et al. 2005)

  - "Amisulpride is an effective 'atypical' antipsychotic drug for those with schizophrenia. Amisulpride may offer a good general profile, at least compared to high-potency 'typical' antipsychotics. It may also yield better results in some specific outcomes related to efficacy, such as improvement of global state and general negative symptoms. It might be more acceptable and more tolerable than high-potency conventional antipsychotics, especially regarding extrapyramidal side-effects."

  - "The large proportion of participants leaving studies early in these trials makes it difficult to draw firm conclusions on olanzapine's clinical effects. For people with schizophrenia it may offer antipsychotic efficacy with fewer extrapyramidal adverse effects than typical drugs, but more weight gain."

Leucht et al. How to read and understand and use systematic reviews and meta-analyses. Acta Psychiatr Scand 2009: 119: 443–450

Publication bias

- *publication bias, file-drawer problem*
- If the results of the study are not desired (e.g. statistically significant), the results are not reported
  - partial solutions: searching for unpublished studies, clinical trials registers
- Either the whole study will be omitted or part of the results will not be published (*outcome reporting bias*)
- Language may affect if a positive result is published more likely in English, otherwise in your own language

Luoto R. Julkaisuharha – Lääketieteellisen tiedon akilleenkantapää. Duodecim 2012; 128: 489-96.

Polyzos NP, Valachis A, Patavoukas E, Papanikolaou EG, Messinis IE, Tarlatzis BC, Devroey P. Publication bias in reproductive medicine: from the European Society of Human Reproduction and Embryology annual meeting to publication. Hum Reprod 2011; 26:1371-6.

BACKGROUND: Treatment decisions should ideally be based on well-designed randomized controlled trials (RCTs). Here we determine the rate of full publication of RCTs presented at annual meetings of the European Society of Human Reproduction and Embryology (ESHRE), identify potential bias against publishing non-significant results and results not favoring the experimental arm, quantify this bias in case it exists, and identify factors associated with time to publication.

METHODS: RCTs presented at ESHRE meetings 2003 and 2004 were recorded. Subsequent search in Medline, Cochrane Library and EMBASE was performed through December 2010 to identify full-text publication in a peer-review journal.

RESULTS: Among 155 abstracts describing RCTs 89 (57%) were published in full-text in a peer-review journal. Median time from presentation to publication was 15 months (range: 0–75). In bivariate analysis, only type of presentation and presence of outcomes favoring the experimental arm were related to publication rate. Studies presented orally or reporting a positive outcome in favor of the experimental arm were more likely to be published ($P = 0.018$ and 0.014, respectively). Results were consistent in a multivariable logistic regression, with odds ratio (OR) 2.51 [95% confidence interval (CI), 1.25–5.03] for oral versus poster presentations and OR 2.46 (95% CI, 1.23–4.95) for trials favoring versus not favoring the experimental arm. Kaplan–Meier curves revealed time to publication was shorter for oral presentations (log-rank test = 0.013) and trials favoring the experimental arm, compared with all others (log rank = 0.007).

CONCLUSIONS: RCTs with significant results in favor of the experimental arm are more likely to be published and are published sooner. Publication bias in reproductive medicine is a fact.

Statistical methods
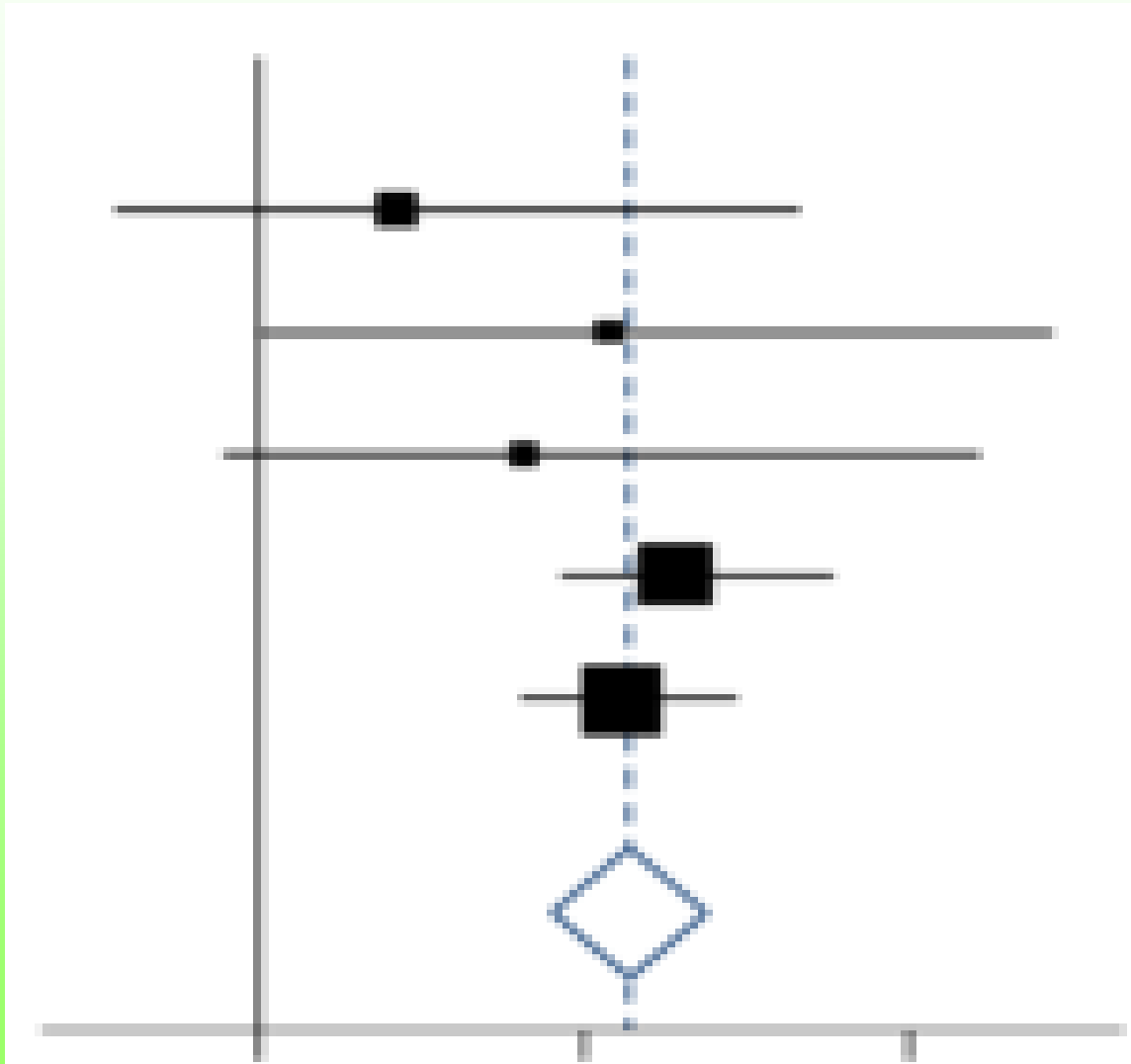
Pooling the studies

- The most important thing when combining studies is the **magnitude and direction of effect** (*effect size*), not statistical significance

- Are the earlier results similar or not?

    - homogeneity/heterogeneity

- Choosing the result (outcome) if duplicate results?

    - sample size, follow-up length, mean of results?

Pooling the studies

- Estimate the effect in each study
- Effect size measures (categorical/continuous variables)
  - odds ratio, relative risk, differences in percentages (absolute risk difference)
  - correlation coefficient (Pearson's r etc.)
  - standardized mean difference
- Unadjusted or adjusted effect
- Sometimes it is only stated that the association was studied, but results were non-significant (N.S.)?
  - first contact authors?
  - if no reply, exclude the studies or set the effect to zero?

Kelley K & Preacher KJ. On Effect Size. Psychological Methods 2012; 17:137-52.

Cohen's d

- Standardized mean difference
- Also Hedges's g, Glass's Δ
- *(group A mean – group B mean) / pooled standard deviation*

$$\text{Cohen's d} = \frac{\overline{X}_1 - \overline{X}_2}{\sigma_{\text{pooled}}}$$

$$\sigma_{\text{pooled}} = \sqrt{\frac{(n_1)\sigma_1^2 + (n_2)\sigma_2^2}{n_1 + n_2}}$$

$$\text{Hedges's g} = \frac{\overline{X}_1 - \overline{X}_2}{s_{\text{pooled}}}$$

$$s_{\text{pooled}} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- **Different effect measures can be estimated from other statistics!**

•Rosenthal & Rubin. Psychological Bulletin 1986;99: 400-6.

•Borenstein ym. Introduction to Meta-Analysis. Wiley, 2009.

	Statistic	d-value
1.	t	$\dfrac{2t}{\sqrt{df}}$
2.	z	$\dfrac{2z}{\sqrt{N}}$
3.	$F\,df_n = 1$	$2\sqrt{\dfrac{F}{df_d}}$
4.	$F\,df_n > 1$	$2\sqrt{\dfrac{df_n F}{df_d}}$
5.	$\chi^2\,df = 1$	$2\sqrt{\dfrac{\chi^2}{N-\chi^2}}$
6.	$\chi^2\,df > 1$	$2\sqrt{\dfrac{\chi^2}{N}}$
7.	r	$\sqrt{\dfrac{4r^2}{1-r^2}}$
8.	d	d

56

Effect size estimates

	Small	Moderate	Large	Very large
Cohen's d	0.2	0.5	0.8	1.3
Pearson's r	0.1	0.3	0.5	0.7
Odds Ratio	1.5	2.5	4	10
Difference in percentage*	7	18	30	45

 * Group difference in percentages, when percentages within 15-85%

Cohen. Psychol Bull 1992;112:155-9; Rosenthal. J Soc Serv Res 1996;21:37-59.

Weighting studies

- Fixed effects method
  - Assumption is that there is one true effect size and that all differences in observed effects are due to sampling error
    - ☞ In observational studies this is in practice unlikely to be true
  - In practice, the weighting is based on the sample size

$$y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \eta_j$$

Weighting studies

- **Random effects method** $\quad y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \ldots + \eta + \varepsilon_j$
  - ◆ Effect is expected to vary between studies and both between-study and within-studies variances are allowed
  - ◆ Weighting is based partly on standard error (sample size), but if heterogeneous results, the weights differ less
  - ◆ If heterogeneity low: fixed effect ≈ random effect
- **Possible useful, but subjective methods?**
  - ◆ Weighting with quality scores
  - ◆ Bayesian meta-analysis

Heterogeneity

- Cochran's Q
  - ◆ Based on the Chi-square test
  - ◆ Depends on number of studies
- Currently recommended to use a variation:

  $I^2 = (Q-df)/Q*100\%$

  - ◆ df = number of studies - 1
  - ◆ 25% low, 50% moderate and 75% high heterogeneity
  - ◆ Interpretation: "what proportion of the total variation across studies is beyond chance"
- If small number of studies both methods lack power

Cochran. Biometrics 1954;10:101-29; Higgins & Thompson. Stat Med 2002;21:1539-58; Ioannidis et al. BMJ 2007;335:914-6.

Meta-regression

- How study characteristics explain heterogeneity in effect size measures?
- Continuous or categorical variables
- Variables are related to studies not to individuals
  - Proportion of men
  - Mean age of the sample
  - Country
  - Research methods (e.g. instruments)
  - Quality of research
  - Study collection period (or year of publication)
  - Patient groups
  - Length of the follow-up

Meta-regression

- Statistical test for the effect of background factor on heterogeneity (t- or z- test)

- Does not adjust the original result

- As an alternative to meta-regression is to examine the direct effect of covariate on the effect

  - Then the research question is different and the data is in a different format (e.g. the effects are collected by gender or data on effect of gender on outcome is collected)

Koskinen J, Löhönen J, Koponen H, Isohanni M, Miettunen J. Prevalence of alcohol use disorders in schizophrenia – a systematic review and meta-analysis.
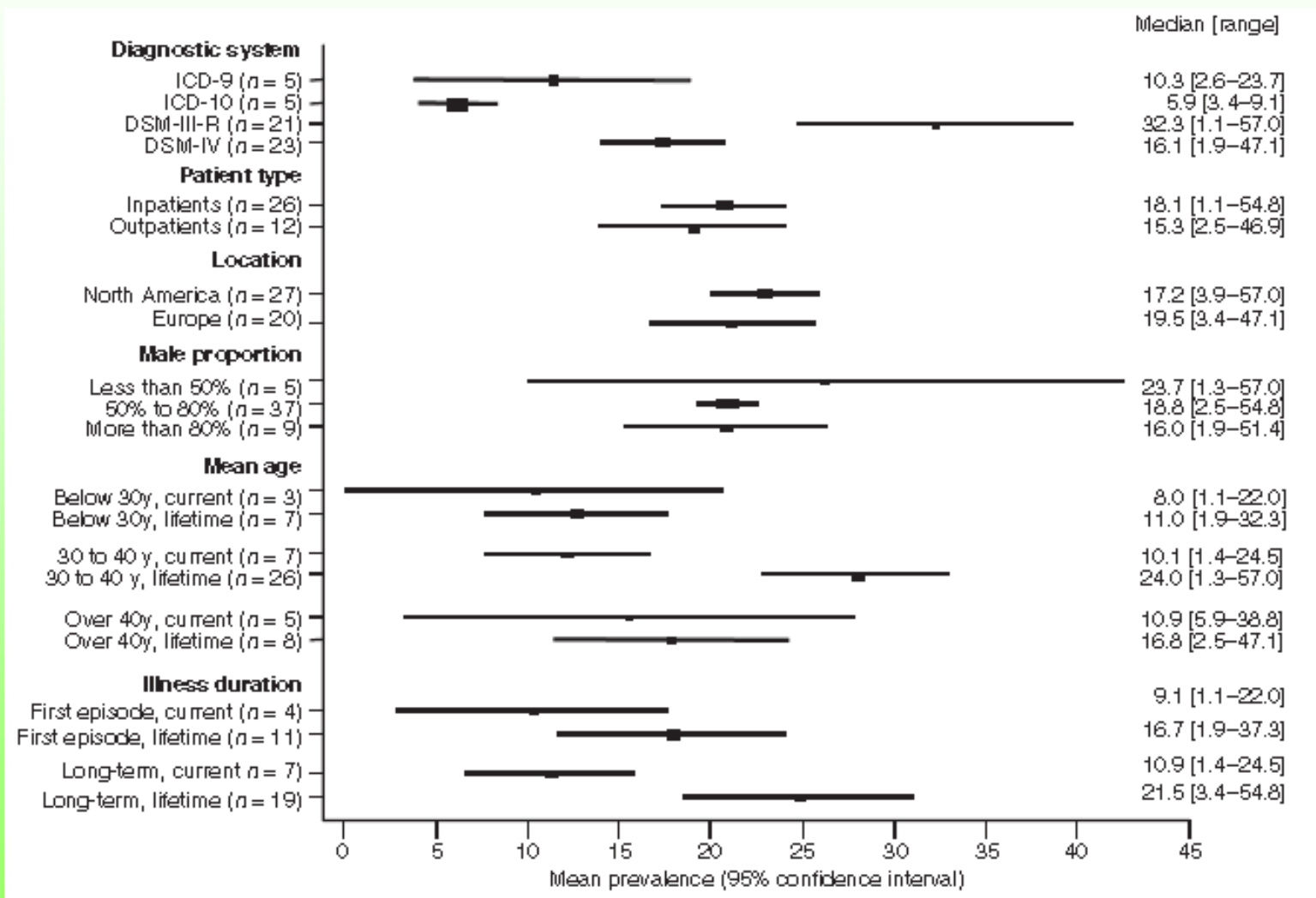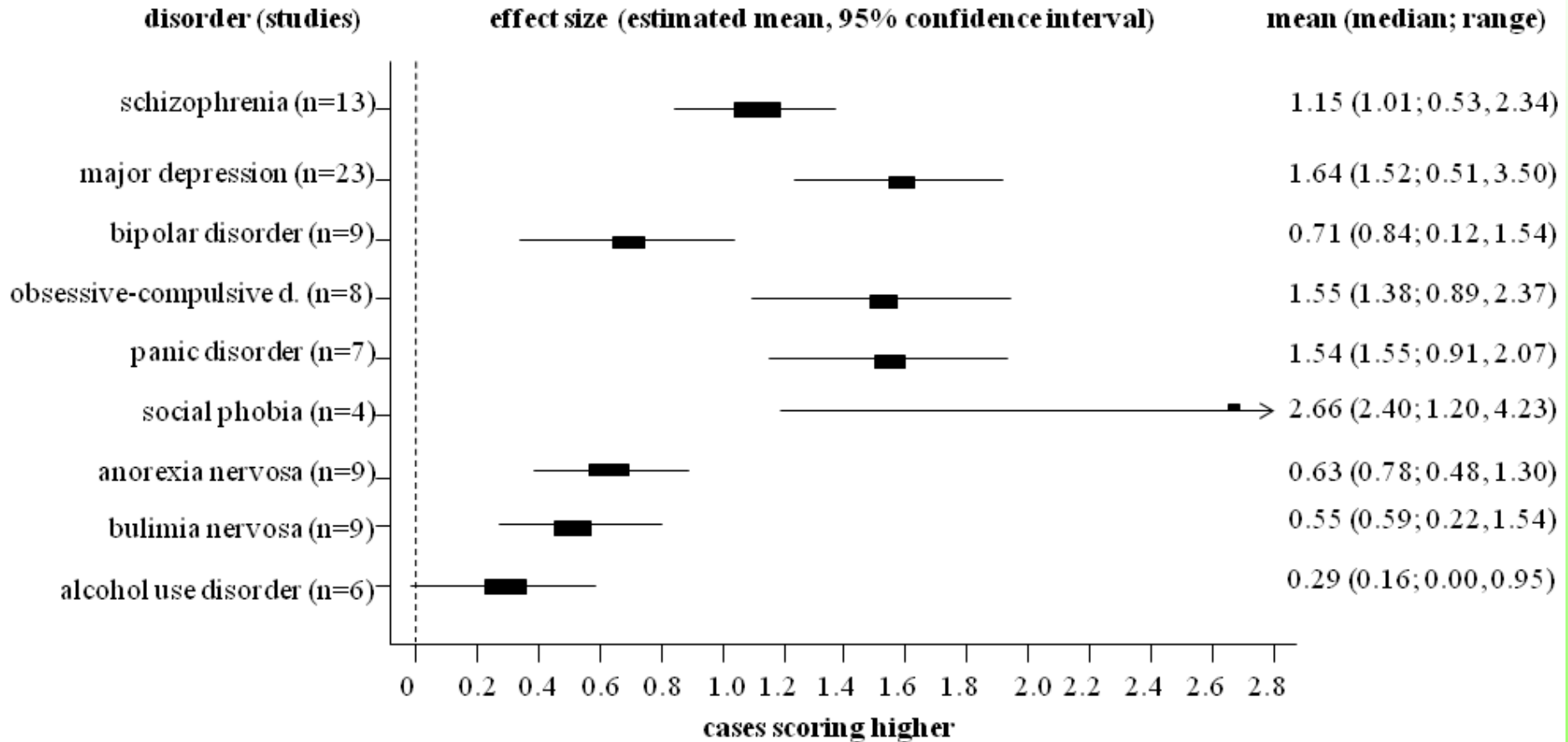
"network meta-analysis"

- To explore a question which has not been directly investigated in original studies

- E.g. combining studies with different drugs compared to placebo, so that the drugs are actually compared to each other

Meta-analysis of temperament traits in case-control studies in different psychiatric disorders

Harm Avoidance

disorder (studies)	effect size (estimated mean, 95% confidence interval)	mean (median; range)
schizophrenia (n=13)		1.15 (1.01; 0.53, 2.34)
major depression (n=23)		1.64 (1.52; 0.51, 3.50)
bipolar disorder (n=9)		0.71 (0.84; 0.12, 1.54)
obsessive-compulsive d. (n=8)		1.55 (1.38; 0.89, 2.37)
panic disorder (n=7)		1.54 (1.55; 0.91, 2.07)
social phobia (n=4)		2.66 (2.40; 1.20, 4.23)
anorexia nervosa (n=9)		0.63 (0.78; 0.48, 1.30)
bulimia nervosa (n=9)		0.55 (0.59; 0.22, 1.54)
alcohol use disorder (n=6)		0.29 (0.16; 0.00, 0.95)

cases scoring higher

Miettunen & Raevuori. Compr Psychiatry 2012

Statistical software

- STATA
  - http://www.biostat.jhsph.edu/~fdominic/teaching/bio656/software/meta.analysis.pdf   (Sterne ym. 2001)
- SPSS
  - http://mason.gmu.edu/~dwilsonb/ma.html
- SAS
  - http://mason.gmu.edu/~dwilsonb/ma.html
- R
  - http://cran.r-project.org/web/views/MetaAnalysis.ht
- Excel, MIX
  - http://mix-for-meta-analysis.info
- Comprehensive Meta-Analysis (CMA), Metawin, RevMan

http://www.meta-analysis.com/pages/comparisons.html

Other statistical methods

- Meta-analyses of diagnostic tests and screens (Hasselblad & Hedges. Psychol Bull 1995;117:167-78)
- Bayesian meta-analysis (Berry. Clin Trials 2009;6:28-41; Schmid. Eval Health Prof 2001;24:165-89; Sutton & Abrams. Stat Meth Med Res 2001; 10: 277–303; Warn et al. Stat Med 2002; 21: 1601-23.
- Meta-analysis of factor analyses (Becker. Psychol Med 1996;1:341-53)
- Meta-analysis of structural equation modeling (Cheung & Chan. Psychol Meth 2005;10:40.64)
- Imputing missing data in meta-analyses of e.g. clinical trials (Higgins et al. Clin Trials 2008;5:225-39)
- etc.

References
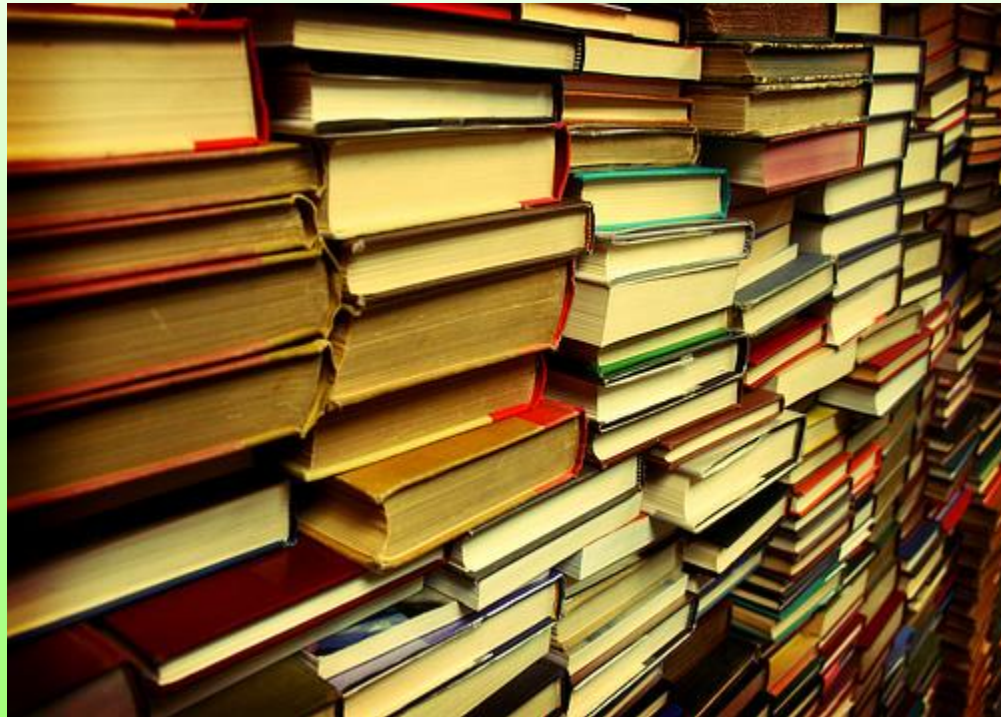
- Blettner M et al. Traditional reviews, meta-analyses, and pooled analyses in epidemiology. Int J Epidemiology 1999; 28: 1-9.

- Borenstein M et al. Introduction to Meta-Analysis. Wiley, 2009.

- Cook DJ et al. (1995) Methodological guidelines for systematic reviews of randomized control trials in health care from the Potsdam consultation on meta-analysis. J Clin Epidemiol 48: 167-71.

- Egger M et al. Systematic reviews in health care. 2. painos. Lontoo: BMJ Publishing Group, 2001.

- Gerber S et al. Bibliographic study showed improving methodology of meta-analyses published in leading journals 1993-2002. J Clin Epidemiol 2007;60:773-80

- Glass GV et al. (eds.). Meta-analysis in Social Research. Beverly Hills, CA: Sage, 1981.

- Guyatt GH et al. (2008) GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336: 924–926.

- Jadad A et al. (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 17:1–12.

- Kelley K & Preacher KJ (2012) On effect size. Psychol Meth 17:137-52.

- Leucht S et al. (2009). How to read and understand and use systematic reviews and meta-analyses. Acta Psychiatr Scand 119: 443–450.

- Liberati A et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6:e1000100.

- Moher D et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097

- Sauerland S & Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. W J Surgery 2005;29:582-7.

- Shea BJ et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Medical Research Methodology 2007; 7:10

- Sterne JAC (ed.). Meta-Analysis in Stata: An Updated Collection from the Stata Journal, 2009.

- Stroup DF et al. Meta-analysis of observational studies in epidemiology. A proposal for reporting. JAMA 2000; 283:2008-12.

- Sutton AJ et al. Methods for Meta-analysis in Medical Research. London: John Wiley, 2000.

KT283