Principles of Systematic Review: Focus on Alcoholism Treatment

Manit Srisurapanont, M.D.
Professor of Psychiatry
Department of Psychiatry,
Faculty of Medicine, Chiang Mai University

For Symposium 1A: Systematic Review of Alcoholism Treatment
2nd National Alcohol Conference, Bangkok, December 13, 2006

Terms (cont.)

- Systematic review[1]: the application of scientific strategies that limit bias by the use of systematic assembly, critical appraisal, and synthesis of all relevant studies on a specific topic
- Meta-analysis[1]: a systematic review that employs statistical methods to combine and summarize the results of several studies

[1]Cook DJ, et al. J Clin Epidemiol 1995;48:167-71.

Narrative review vs. systematic review[1]

Features	Narrative review	Systematic review
Question	Often broad in scope	Often a focused clinical question
Sources and search	Not usually specified, potentially biased	Comprehensive sources and explicit search strategy
Selection	Not usually specified, potentially biased	Criterion-based selection, uniformly applied
Appraisal	Variable	Rigorous critical appraisal
Synthesis	Often a qualitative summary	Quantitative summary, e.g., meta-analysis
Inference	Sometimes evidence-based	Usually evidence-based

[1]Cook DJ, et al. Ann Intern Med 1997;126:376-80.

Level of evidence[1]

1. Meta-analysis, systematic reviews of RCTs, or RCTs
2. Systematic reviews of case-control or cohort studies
3. Case-control or cohort studies
4. Non-analytic studies, e.g., case reports, case series
5. Expert opinions

[1]Modified from Harbour R, et al. BMJ 2001;323:334-6.

Steps in conducting SR

1. Defining objectives and review questions
2. Defining study selection criteria
3. Search for trials
4. Selection of trials
5. Study quality assessment
6. Data extraction
7. Data synthesis (+/- sensitivity analysis)
8. Publication bias

(9. Interpretation and discussion)

Objectives/review questions

- Components include
  - Population or participants
  - Interventions: usually a comparison between two or more alternatives
  - Outcomes: clinical +/- economic
  - Study designs
- With more details, these can be use for defining the study selection criteria

Search for & selection of trials

- Sources of research evidence:
  - Electronic bibliographic database, e.g., MEDLINE, EMBASE, Cochrane Controlled Trials Register (CCTR), CINAHL, LILAC
  - Reference lists from relevant primary and review articles, hand searching, grey literature, and conference proceedings
  - Research registers, researcher, and manufacturers
- Selection of trials in an unbiased way and based on selection criteria

Quality assessment & data extraction

- Assess the study quality by using:
  - Individual quality components or items
  - Quality checklists
  - Quality scales
- Data extraction: dealing with human error and missing data

Data synthesis: general

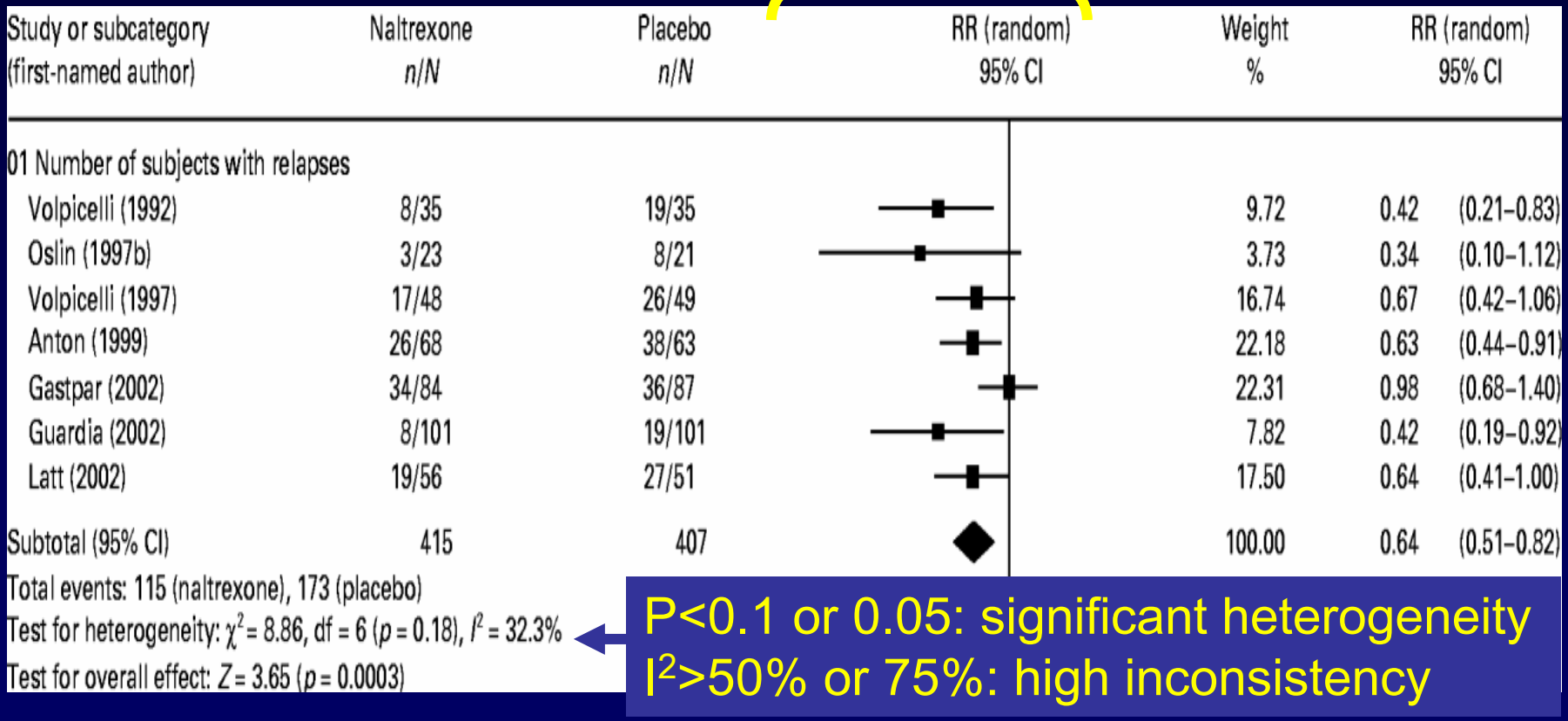
- When data are too sparse, of too low quality, or too heterogeneous to proceed with their statistical aggregation, perform a narrative, descriptive (qualitative) summary with/without graphs and tables and AVOID meta-analysis
- Include all relevant and clinically meaningful measures of treatment effect, especially, both risks and benefits
- Intention-to-treat

Data synthesis: heterogeneity

- Test for heterogeneity: some meta-analysis software, e.g., RevMan, can automatically compute the magnitude of heterogeneity (p-value) by using Chi-square and I-square tests
- Suggested steps in exploring for heterogeneity of results
  1. Graphical exploration, e.g., forest plot
  2. Statistical tests of heterogeneity
  3. Subgroup analysis
  4. Statistical regression modeling

Data synthesis: heterogeneity (cont.)

Forest plot



Study or subcategory (first-named author)	Naltrexone n/N	Placebo n/N	RR (random) 95% CI	Weight %	RR (random) 95% CI
01 Number of subjects with relapses					
Volpicelli (1992)	8/35	19/35		9.72	0.42 (0.21–0.83)
Oslin (1997b)	3/23	8/21		3.73	0.34 (0.10–1.12)
Volpicelli (1997)	17/48	26/49		16.74	0.67 (0.42–1.06)
Anton (1999)	26/68	38/63		22.18	0.63 (0.44–0.91)
Gastpar (2002)	34/84	36/87		22.31	0.98 (0.68–1.40)
Guardia (2002)	8/101	19/101		7.82	0.42 (0.19–0.92)
Latt (2002)	19/56	27/51		17.50	0.64 (0.41–1.00)
Subtotal (95% CI)	415	407		100.00	0.64 (0.51–0.82)
Total events: 115 (naltrexone), 173 (placebo)					
Test for heterogeneity: $\chi^2 = 8.86$, df = 6 ($p = 0.18$), $I^2 = 32.3\%$					
Test for overall effect: $Z = 3.65$ ($p = 0.0003$)					

P<0.1 or 0.05: significant heterogeneity
I^2>50% or 75%: high inconsistency

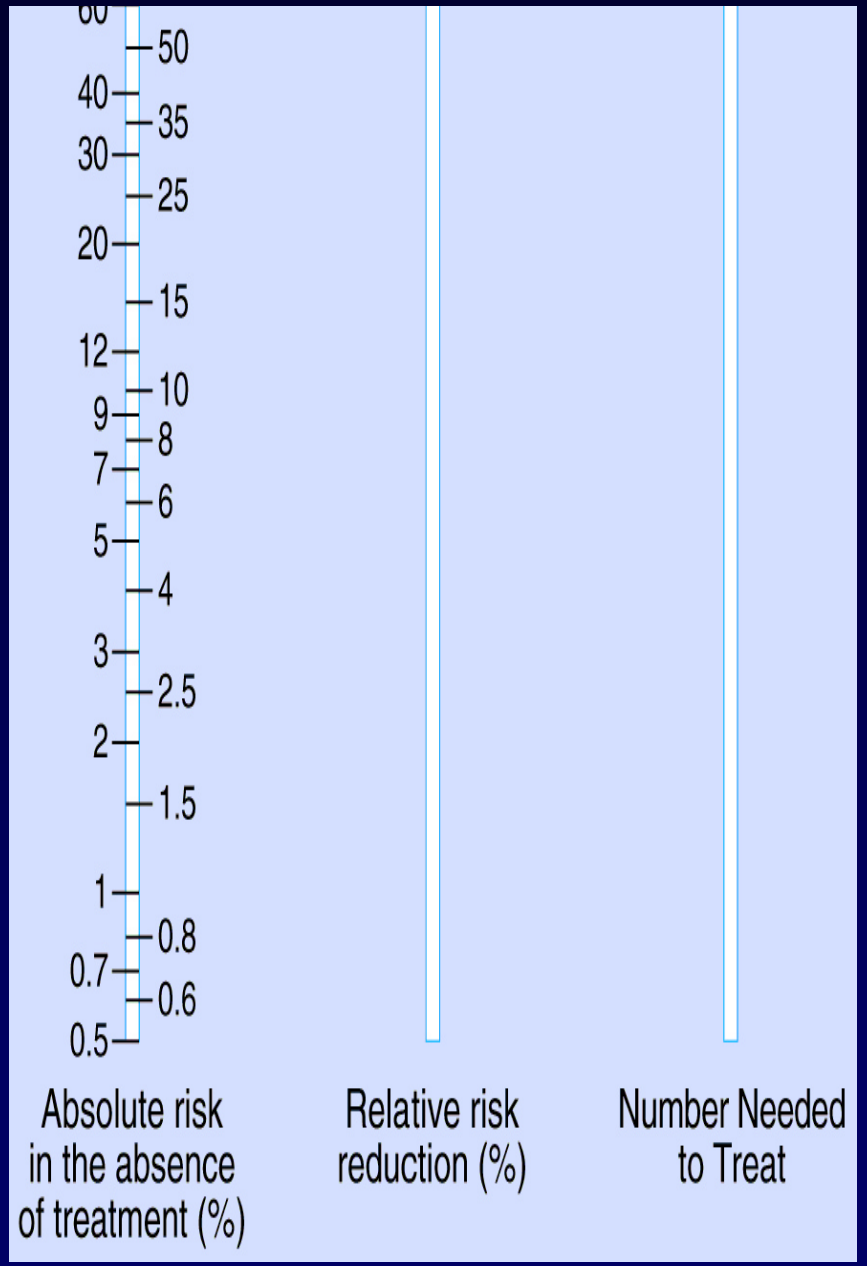
From: Srisurapanont M, Jarusuraisin N. Int J Neuropsychopharmacol 2005;8:267-80.

Data synthesis: heterogeneity (cont.)

- If heterogeneity exists:
  - avoid the use of a fixed effect model
  - examine potential sources of heterogeneity (e.g., differences in study quality, participants, interventions, or in the definitions and measures of outcomes)
- The interpretation of statistical evidence of heterogeneity, as well as what to do when heterogeneity is present, are matters still to be settled

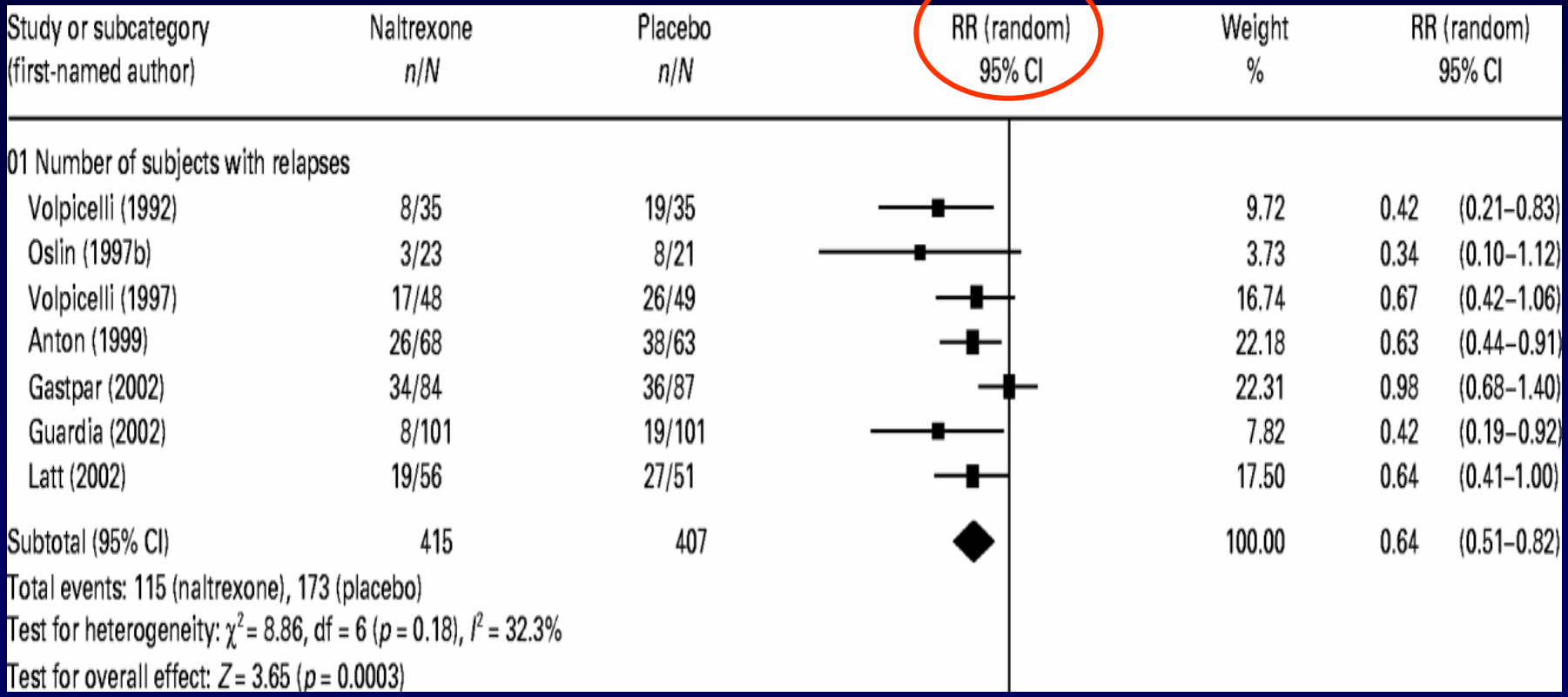
Data synthesis: dichotomous data

- Dichotomous data, e.g., odds ratios (ORs), relative risks (RRs), absolute risks (ARs), number needed to treat (NNT)



Data synthesis: dichotomous data (cont.)

Relative Risk



Study or subcategory (first-named author)	Naltrexone n/N	Placebo n/N	RR (random) 95% CI	Weight %	RR (random) 95% CI
01 Number of subjects with relapses					
Volpicelli (1992)	8/35	19/35		9.72	0.42 (0.21–0.83)
Oslin (1997b)	3/23	8/21		3.73	0.34 (0.10–1.12)
Volpicelli (1997)	17/48	26/49		16.74	0.67 (0.42–1.06)
Anton (1999)	26/68	38/63		22.18	0.63 (0.44–0.91)
Gastpar (2002)	34/84	36/87		22.31	0.98 (0.68–1.40)
Guardia (2002)	8/101	19/101		7.82	0.42 (0.19–0.92)
Latt (2002)	19/56	27/51		17.50	0.64 (0.41–1.00)
Subtotal (95% CI)	415	407		100.00	0.64 (0.51–0.82)
Total events: 115 (naltrexone), 173 (placebo)					
Test for heterogeneity: $\chi^2 = 8.86$, df = 6 ($p = 0.18$), $I^2 = 32.3\%$					
Test for overall effect: $Z = 3.65$ ($p = 0.0003$)					

From: Srisurapanont M, Jarusuraisin N. Int J Neuropsychopharmacol 2005;8:267-80.

Data synthesis: continuous data

- Continuous data
  - Weighted mean difference (WMD): for an outcome assessed by the same scale (e.g., weight)
  - Standardized mean difference (SMD): for an outcome assessed by different scales (e.g., pain)

Data synthesis: continuous data (cont.)

Outcome comparison (studies)	Total number of patients	Test for heterogeneity	Mean effect size (95% CI)
Naltrexone vs. placebo (short-term secondary outcomes only)			
Time to first drink (Anton et al., 1999; Guardia et al., 2002; Hersh et al., 1998; Kranzler et al., 2000)	511	$I^2 = 0\%$	WMD −0.06 (−1.04 to 0.93)
Drinking days (Hersh et al., 1998; Kranzler et al., 2000; Latt et al., 2002; O'Malley et al., 1992; Volpicelli et al., 1997)	489	$I^2 = 61.5\%$	WMD −1.96 (−5.47 to 1.56)
Standard drinks (Anton et al., 1999; Chick et al., 2000; Guardia et al., 2002; Hersh et al., 1998; Kranzler et al., 2000; O'Malley et al., 1992)	772	$I^2 = 65.3\%$	SMD −0.21 (−0.46 to 0.04)
Craving (Anton et al., 1999; Kranzler et al., 2000; O'Malley et al., 1992; Volpicelli et al., 1997)	400	$I^2 = 34.1\%$	SMD −0.10 (−0.35, 0.15)

From: Srisurapanont M, Jarusuraisin N. Int J Neuropsychopharmacol 2005;8:267-80.

Data synthesis: effect size

- Continuous data (cont.)

  effect size (*d*)[1]

  $$d = \frac{\text{mean (exp. group)} - \text{mean (control group)}}{\text{pooled standard deviation}}$$

  d [2] = small = 0.2

  = medium = 0.5 (an effect likely to be visible to the naked eye of a careful observer)

  = large = 0.8

[1]Hedges LV & Olkin I. Statistical methods for meta-analysis, 1985.
[2]Cohen J. A power primer. Psychol Bull 1992,112:155-9.

Data synthesis: effect models

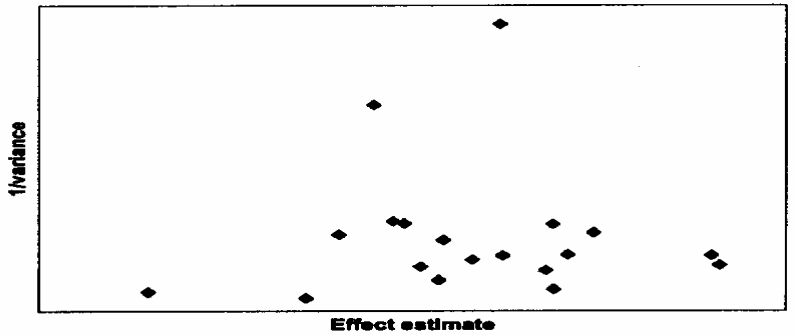
- Fixed effect model: a mathematical model for combining the results of studies that assumes that the effect is truly constant in all the populations studies (homogenous data)
- Random effect model: a mathematical model for combining the results of studies that allows for variation in the effect amongst the populations studies (heterogenous data)
- If the data are perfectly homogenous, the use of random effect model will lead to the same results as that of fixed effect model

Publication bias

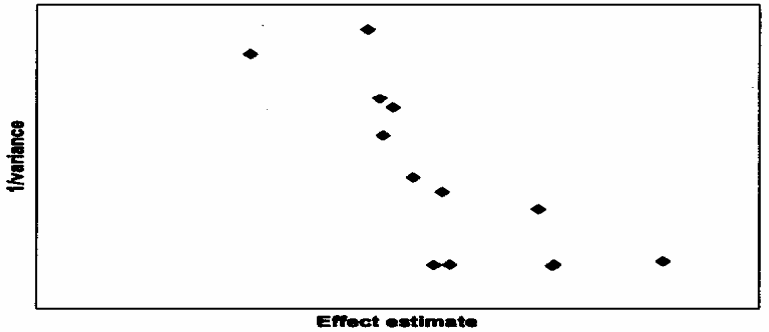
Examples of funnel plots and reasons for asymmetry

Examples of funnel plots

Symmetrical



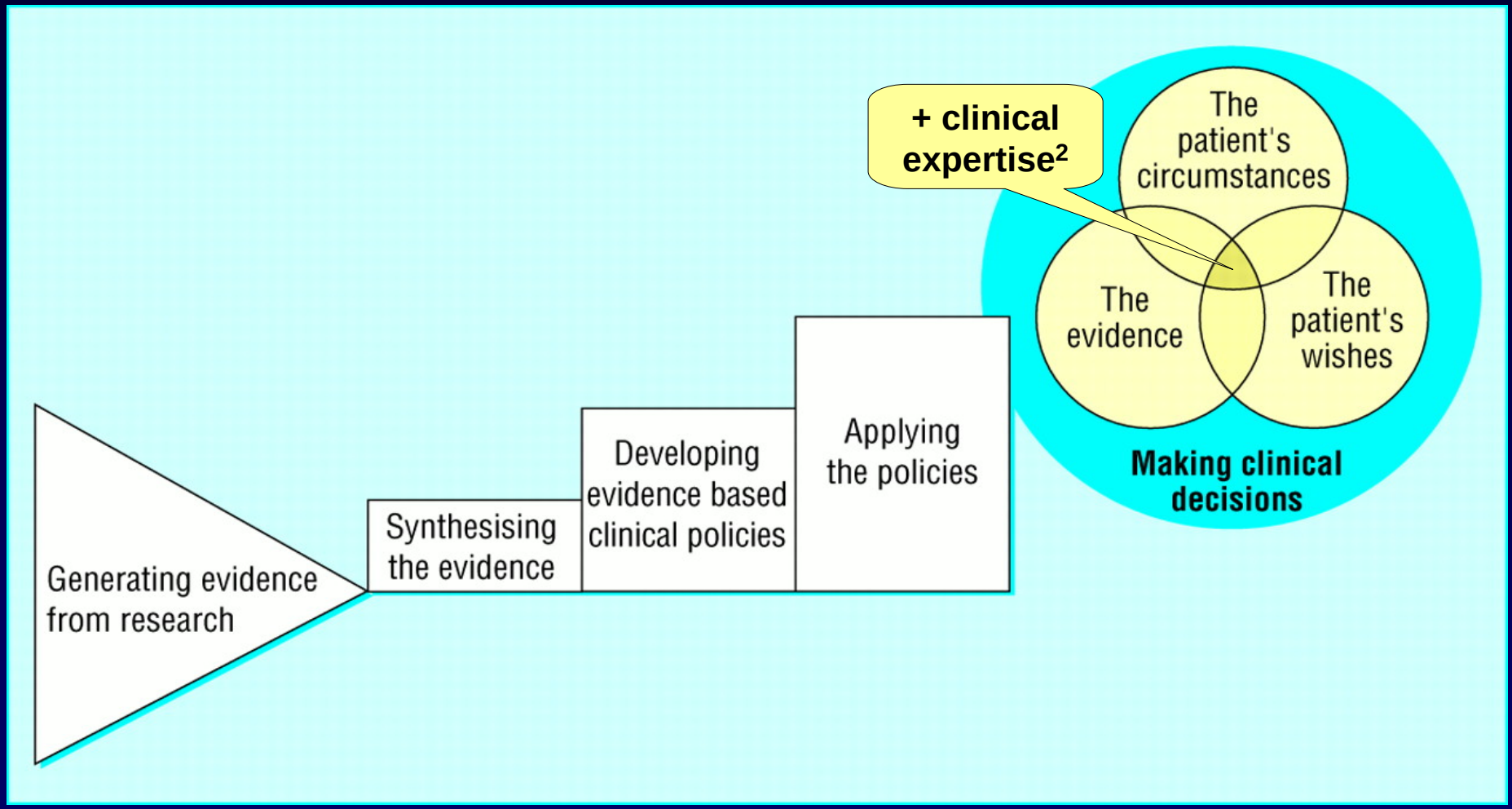
Asymmetrical



Reasons for funnel plot asymmetry

- Publication bias
- Location biases
- English language bias
- Database bias
- Citation bias
- Multiple publication bias
- Bias in provision of data
- Poor methodological quality of small studies
- Clinical heterogeneity e.g. small studies in high risk populations

From evidence to practice[1]



[1]Haynes RB, Haines A. BMJ 1998;317:273-6.
[2]Haynes RB, et al. EBM 2002;7:36-8.