Combining studies: from heterogeneity to similarity

Jos Verbeek, Jani Ruotsalainen
Cochrane OSH Review Group
Finnish Institute of Occupational Health
Kuopio Finland

Finnish Institute of Occupational Health

A systematic review

1. Well-formulated question (PICO)
2. Thorough search
3. Objective selection of studies
4. Critical assessment of methodological quality
5. Objective data extraction
6. Synthesis of the data
   a) appropriate comparisons of interventions and controls
   b) meta-analysis per comparison
7. Conclusions for practice and research

Finnish Institute of Occupational Health

Scope of a review

- Interventions for.....
  - Cochrane Library
    - intervention* NOT pharmacological in title
      - 411 reviews and 261 protocols
- Resulting in a variety of interventions
  - non-drug
  - complex interventions
    - multi-faceted
    - multi-component
    - behavioural
    - team based
    - community-based
    - rehabilitation
    - exercises for..
    - educational

Finnish Institute of Occupational Health

When to combine studies?

- Usual text in Cochrane Protocol Method Section
  - We will assess clinical heterogeneity by examining types of participants, interventions, and outcomes in each study
  - We will pool data from studies judged to be clinically homogeneous with RevMan 5 software.

- When you are going to look for heterogeneity you will probably find it and then it will be difficult to pool studies

- Better to state: We will look for *similarity* between studies

Finnish Institute of Occupational Health

Heterogeneity? Similarity?

- In the review “Interventions for preventing noise-induced hearing loss in workers” you find the following 8 studies. How are they similar assuming that study design is the same in all?

1. Ear muffs vs plugs for preventing hearing loss (STS) in construction workers
2. Legislation versus no legislation for reducing noise exposure (dB) in mines
3. Worker training in ear plug use vs no training for reducing noise exposure (dB) in metal sheet workers
4. Subsidies for employers vs no subsidies for reducing noise exposure (dB) in offshore oil platforms
5. Stronger ear plugs vs lighter ear plugs for preventing hearing loss (STS) in farmers
6. Magnesium vs placebo for preventing hearing loss (STS) in noise-exposed workers
7. Inspections and penalties for preventing noise exposure (dB) in construction industry
8. Occupational health advice to decrease noise exposure (dB) in small companies

Finnish Institute of Occupational Health

Heterogeneity? Similarity?

- In the review “Interventions for preventing noise-induced hearing loss in workers” you find the following 8 studies. How are they similar assuming that study designs are all similar?

1. *Hearing Protection for workers*
- Ear muffs vs plugs for preventing hearing loss (STS) in construction workers
- Stronger ear plugs vs lighter ear plugs for preventing hearing loss (STS) in farmers

- Worker training in ear plug use vs no training for reducing noise exposure (dB) in metal sheet workers

2. *Incentives for employers/firms*
- Legislation versus no legislation for reducing noise exposure (dB) in mines
- Subsidies for employers vs no subsidies for reducing noise exposure (dB) in offshore oil platforms
- Inspections and penalties for preventing noise exposure (dB) in construction industry
- Occupational health advice to decrease noise exposure (dB) in small companies

3. *Drugs to prevent hearing loss in workers*
- Magnesium vs placebo for preventing hearing loss (STS) in noise-exposed workers

Finnish Institute of Occupational Health

Too heterogeneous for meta-analysis?

- Cochrane Systematic Review, Rehabilitation for older people in long-term care, CD004294
- Objective: to evaluate physical rehabilitation interventions directed at improving physical function among older people in long-term care.
- ...From these, 49 studies fulfilled the eligibility criteria and are included in this review.
- ...The included studies are heterogeneous. They examine different types of intervention, and evaluate them with a wide battery of outcome measures. *Such variety made a meta-analysis unfeasible.*
- *Are the authors correct?*

Finnish Institute of Occupational Health

Solutions for heterogeneity 1: focus

1. Narrow down the scope of the review

2. When few studies expected, formulate on beforehand which comparisons will be judged sufficiently similar to be combined.

Finnish Institute of Occupational Health

Need for intervention classification

- Preferably classification should…
  - be based on mechanism
  - have a practical meaning
  - have consequences in resource use

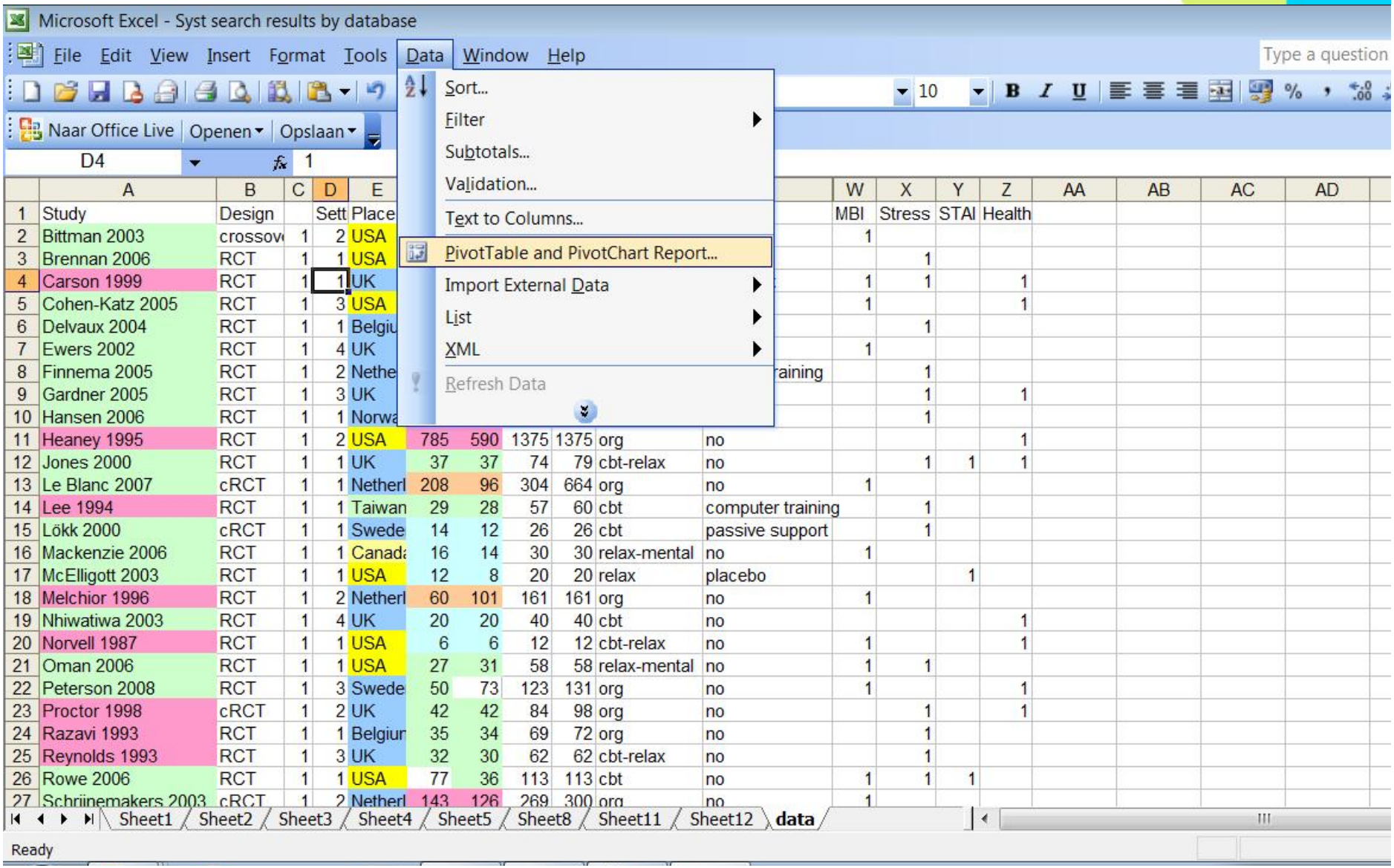
Finnish Institute of Occupational Health

Criteria for intervention classification

- Outcome
  - that the intervention aims to reduce
  - exposure, worker behaviour, occupational disease, disability, injury
- Mode of action
  - environmental, behavioural, clinical
- Level or point of action
  - individual, group, societal level (legal)
- Complexity
  - simple, multi-component, multi-actor
- Target Group
  - workers, students, specific occupations
- Place of delivery or setting
  - hospital, primary care, workplace
- Moment of application
  - preventive (without request for help), treatment
- Mode of delivery
  - verbal, written, web-based, media

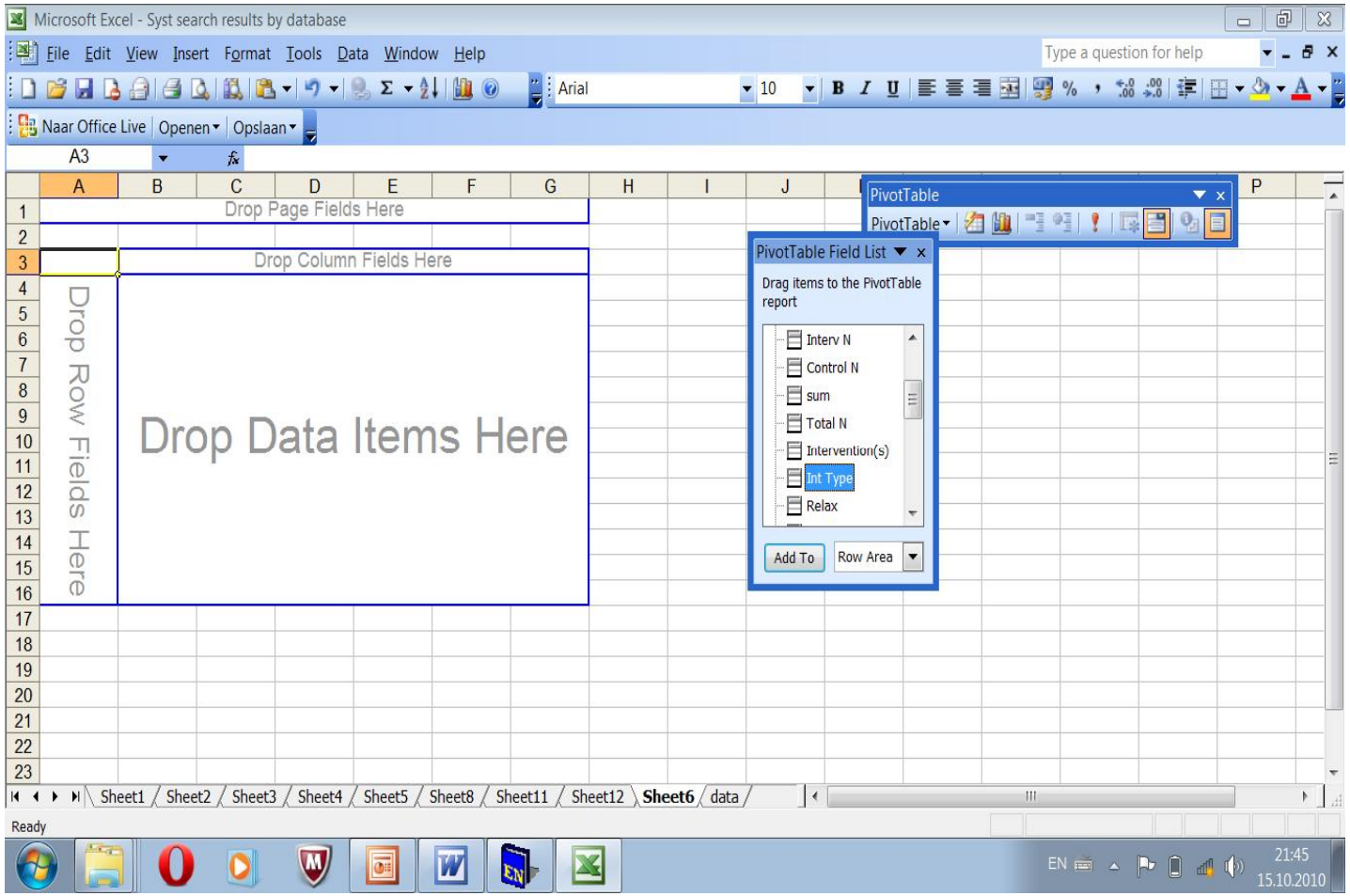
Finnish Institute of Occupational Health

Use Excel Pivot Table



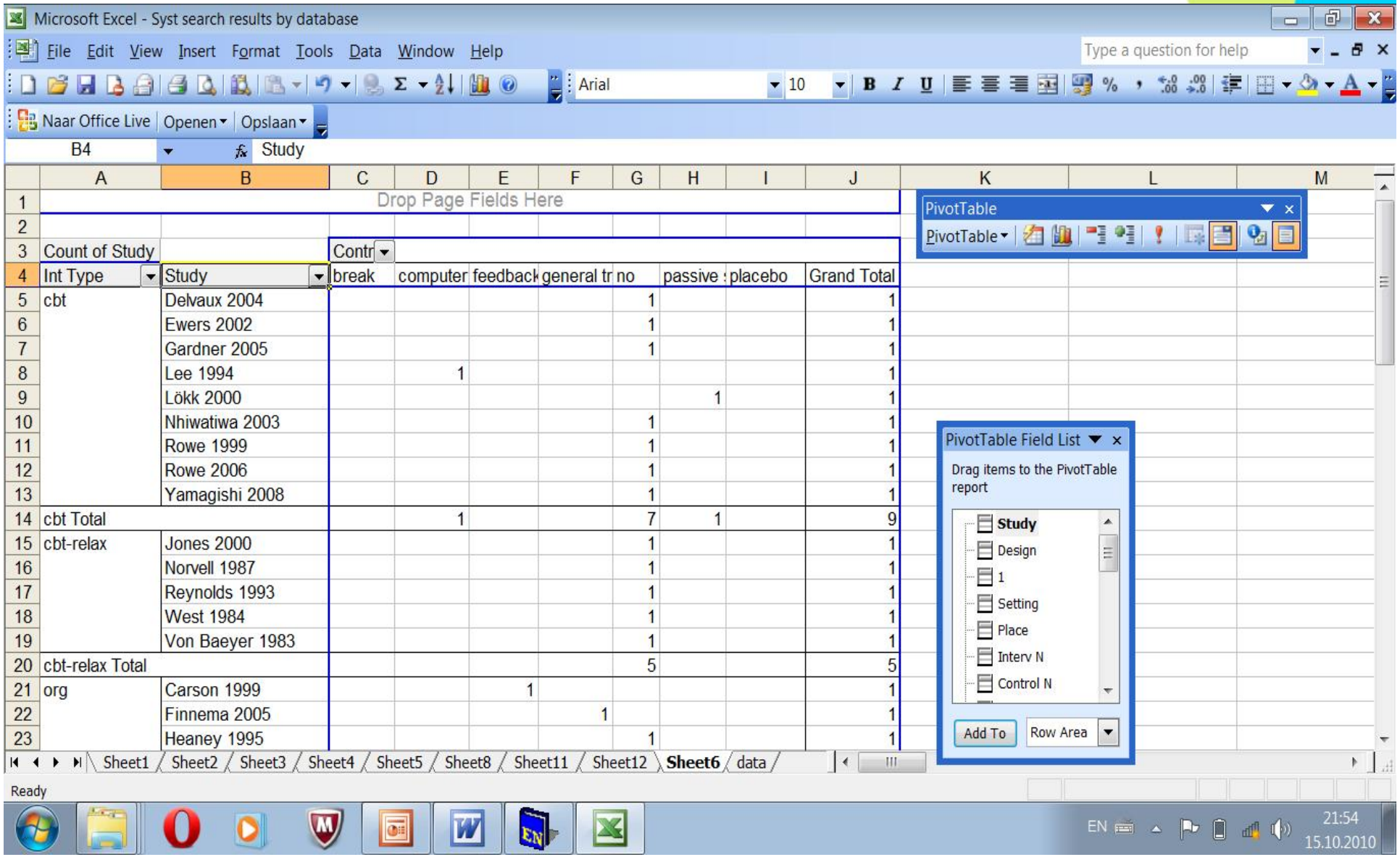
	A	B	C	D	E	W	X	Y	Z	AA	AB	AC	AD
1	Study	Design		Sett	Place	MBI	Stress	STAI	Health				
2	Bittman 2003	crossov	1	2	USA	1							
3	Brennan 2006	RCT	1	1	USA		1						
4	Carson 1999	RCT	1	1	UK	1	1		1				
5	Cohen-Katz 2005	RCT	1	3	USA	1			1				
6	Delvaux 2004	RCT	1	1	Belgiu		1						
7	Ewers 2002	RCT	1	4	UK	1							
8	Finnema 2005	RCT	1	2	Nethe	raining		1					
9	Gardner 2005	RCT	1	3	UK			1	1				
10	Hansen 2006	RCT	1	1	Norwa			1					
11	Heaney 1995	RCT	1	2	USA	785 590 1375 1375 org no			1				
12	Jones 2000	RCT	1	1	UK	37 37 74 79 cbt-relax no		1	1	1			
13	Le Blanc 2007	cRCT	1	1	Netherl	208 96 304 664 org no	1						
14	Lee 1994	RCT	1	1	Taiwan	29 28 57 60 cbt computer training		1					
15	Lökk 2000	cRCT	1	1	Swede	14 12 26 26 cbt passive support		1					
16	Mackenzie 2006	RCT	1	1	Canada	16 14 30 30 relax-mental no	1						
17	McElligott 2003	RCT	1	1	USA	12 8 20 20 relax placebo			1				
18	Melchior 1996	RCT	1	2	Netherl	60 101 161 161 org no	1						
19	Nhiwatiwa 2003	RCT	1	4	UK	20 20 40 40 cbt no				1			
20	Norvell 1987	RCT	1	1	USA	6 6 12 12 cbt-relax no	1			1			
21	Oman 2006	RCT	1	1	USA	27 31 58 58 relax-mental no	1	1					
22	Peterson 2008	RCT	1	3	Swede	50 73 123 131 org no	1			1			
23	Proctor 1998	cRCT	1	2	UK	42 42 84 98 org no		1		1			
24	Razavi 1993	RCT	1	1	Belgiur	35 34 69 72 org no		1					
25	Reynolds 1993	RCT	1	3	UK	32 30 62 62 cbt-relax no		1					
26	Rowe 2006	RCT	1	1	USA	77 36 113 113 cbt no	1	1	1				
27	Schrijnemakers 2003	cRCT	1	2	Netherl	143 126 269 300 org no	1						

Excel Pivot Table



Finnish Institute of Occupational Health

Pivot Table



	A	B	C	D	E	F	G	H	I	J
1	Drop Page Fields Here									
2										
3	Count of Study		Contr							
4	Int Type	Study	break	computer	feedback	general tr	no	passive	placebo	Grand Total
5	cbt	Delvaux 2004					1			1
6		Ewers 2002					1			1
7		Gardner 2005					1			1
8		Lee 1994		1						1
9		Lökk 2000						1		1
10		Nhiwatiwa 2003					1			1
11		Rowe 1999					1			1
12		Rowe 2006					1			1
13		Yamagishi 2008					1			1
14	cbt Total			1			7	1		9
15	cbt-relax	Jones 2000					1			1
16		Norvell 1987					1			1
17		Reynolds 1993					1			1
18		West 1984					1			1
19		Von Baeyer 1983					1			1
20	cbt-relax Total						5			5
21	org	Carson 1999			1					1
22		Finnema 2005				1				1
23		Heaney 1995					1			1

Alternatives for dealing with heterogeneity?

- Because the studies were too heterogeneous to combine them in a meta-analysis we used a levels of evidence synthesis (best evidence synthesis) to combine them

- Data synthesis: The selected studies were very heterogeneous in types of interventions, types of complaints, study population and outcomes measures, and therefore meta-analyses were not performed. Findings were reported narratively.

- *Levels of evidence:* For a more qualitative approach to synthesise the findings from included studies, so-called 'levels of evidence' were used (Ostelo 2002; Van Tulder 1997; Van Tulder 2001).

- Levels of evidence:
  1. Good evidence - provided by generally consistent findings in two or more high-quality studies
  2. Moderate evidence - provided by generally consistent findings in one high-quality study and one or more low-quality studies, or by generally consistent findings in two or more low-quality studies
  3. Limited or conflicting evidence - only one study (either high or low quality), or inconsistent findings in two or more studies
  4. No evidence - no studies. (Henken 2007)

- Are the authors correct?

Finnish Institute of Occupational Health

Too heterogeneous to combine thus..

- Worker training to prevent injuries

- Outcome: Reported Injuries
  - Peterson 2001 found after one year follow-up:
    - intervention: 15 injuries / 450 workers
    - control: 19 injuries / 370 workers
    - RR 0.61 (95% CI 0.3 to 1.2)
    - author's conclusions: non-significant outcome

  - Hansson 2004 found after one year follow-up:
    - intervention: 15 injuries / 402.000 working hours
    - control 23 injuries / 386.000 working hours
    - RR 0.63 (95% CI 0.3 to 1.2)
    - author's conclusions: non-significant outcome

- No quantitative analysis possible we combined studies qualitatively:
  - conclusion: based on two studies with a non-significant outcome we found no evidence of effectiveness

Finnish Institute of Occupational Health

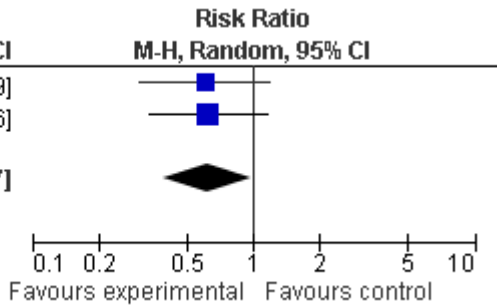
Solution to heterogeneity 2: recalculate

- Recalculate all outcomes on similar scale
  - 2000 working hours = 1 working year (US)

- Combine in meta-analysis

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
Hansson 2004	14	450	19	370	45.6%	0.61 [0.31, 1.19]
Peterson 2001	15	200	23	192	54.4%	0.63 [0.34, 1.16]
Total (95% CI)		**650**		**562**	**100.0%**	**0.62 [0.39, 0.97]**
Total events	29		42			

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 0.00$, df = 1 (P = 0.94); $I^2 = 0\%$
Test for overall effect: Z = 2.07 (P = 0.04)



- Review Conclusion:
  - the intervention reduces injuries with 38%

Finnish Institute of Occupational Health

Solution to heterogeneity 3: narrative synthesis

- Studies used different outcomes and interventions and therefore we did not combine them but described them in a narrative way.

- How would you perform a narrative synthesis?

Finnish Institute of Occupational Health

Narrative synthesis

- Rodgers et al 2009:
  - developing a theory of how the intervention works, why and for whom;
  - developing a preliminary synthesis;
  - exploring relationships within and between studies;
  - assessing the robustness of the synthesis product.

Finnish Institute of Occupational Health

Narrative synthesis

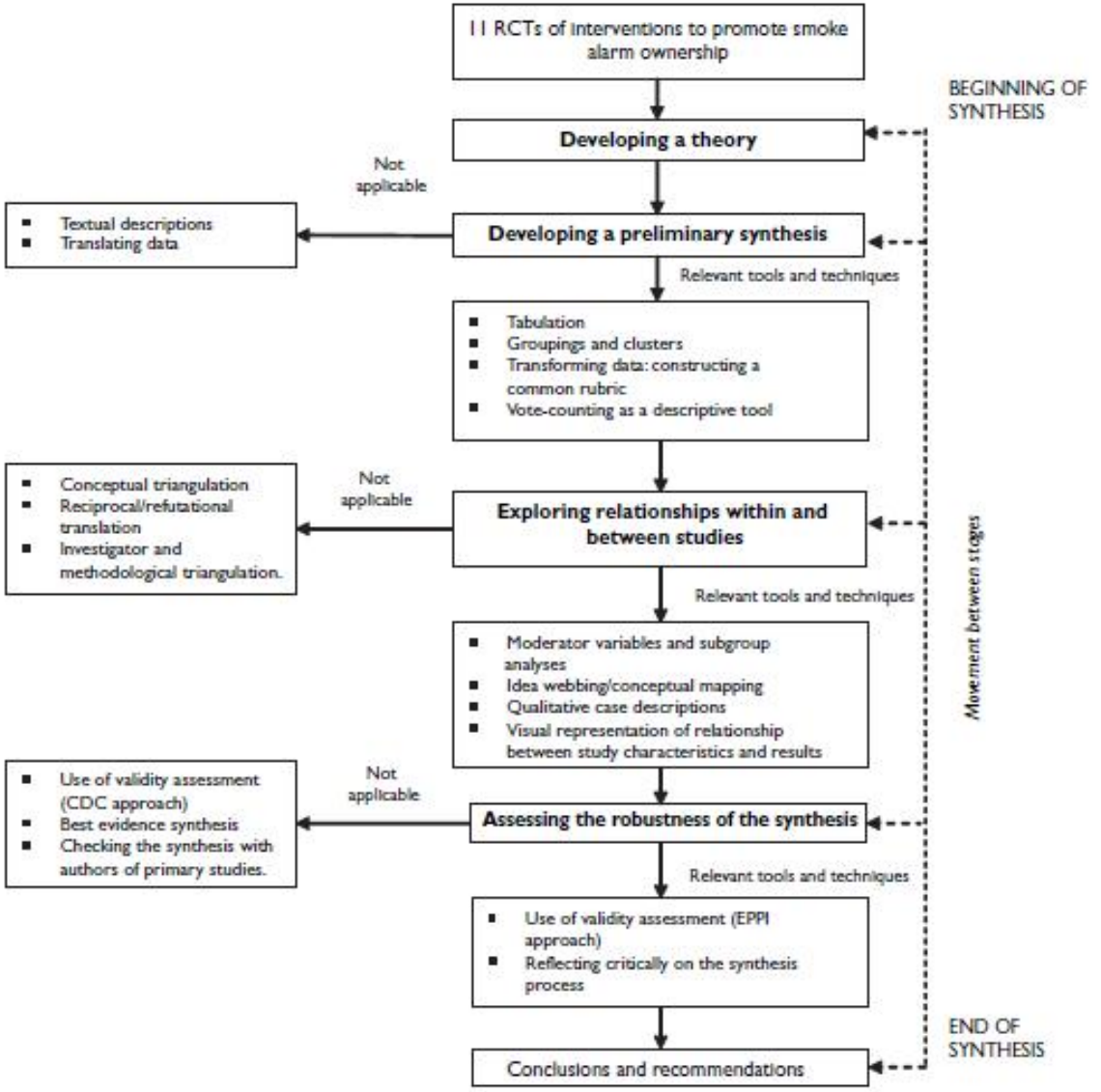


Figure 1. Synthesis Process

Finnish I
Occupati

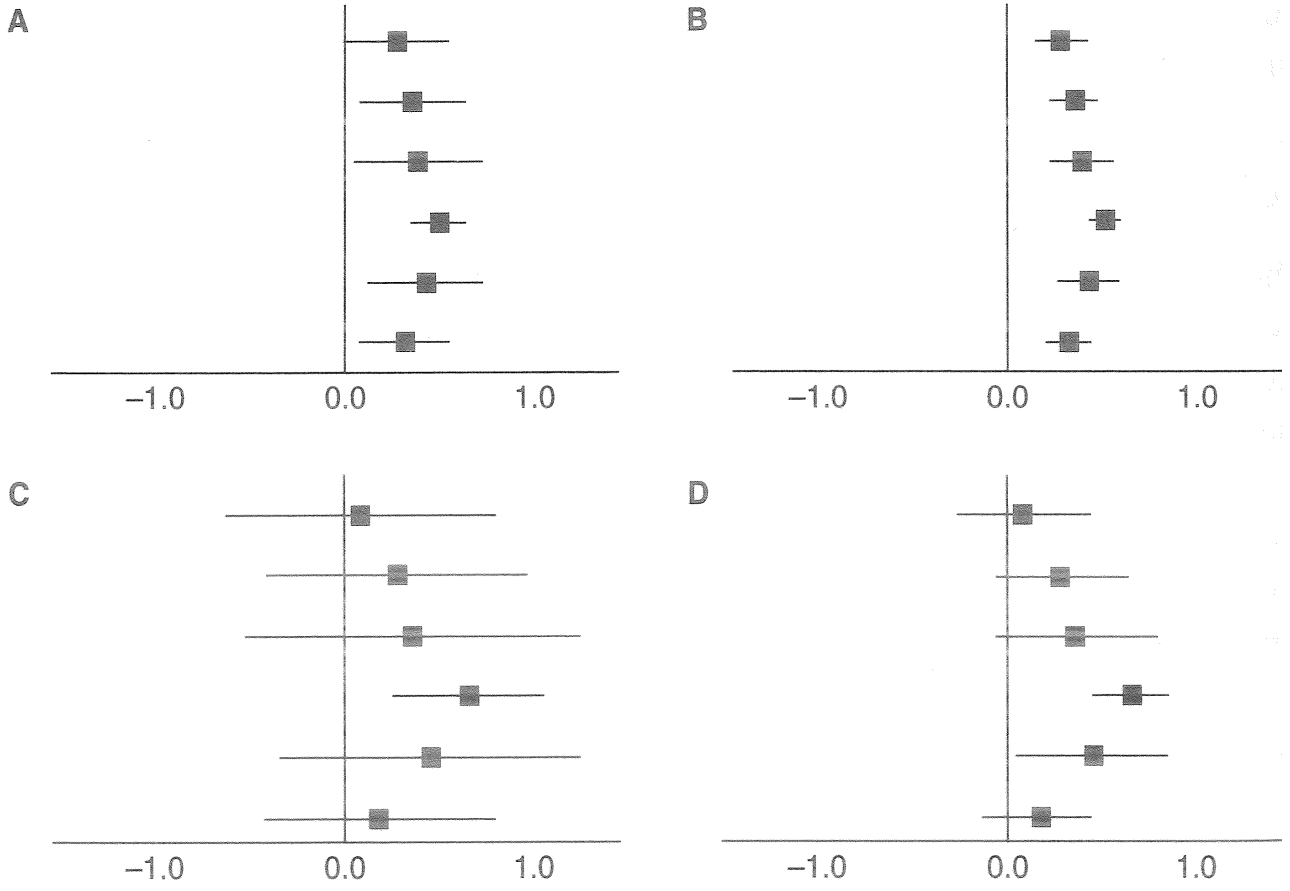
Narrative synthesis

- Rodgers et al. Evaluation 2009 (15) 47-79

- Guidance-led narrative synthesis against a meta-analysis of the same study data.
  - The conclusions of the two syntheses were broadly similar.
  - However, conclusions about the
    - impact of moderators of effect appeared stronger when derived from the meta-analysis,
    - whereas implications for future research appeared more extensive when derived from the narrative synthesis.

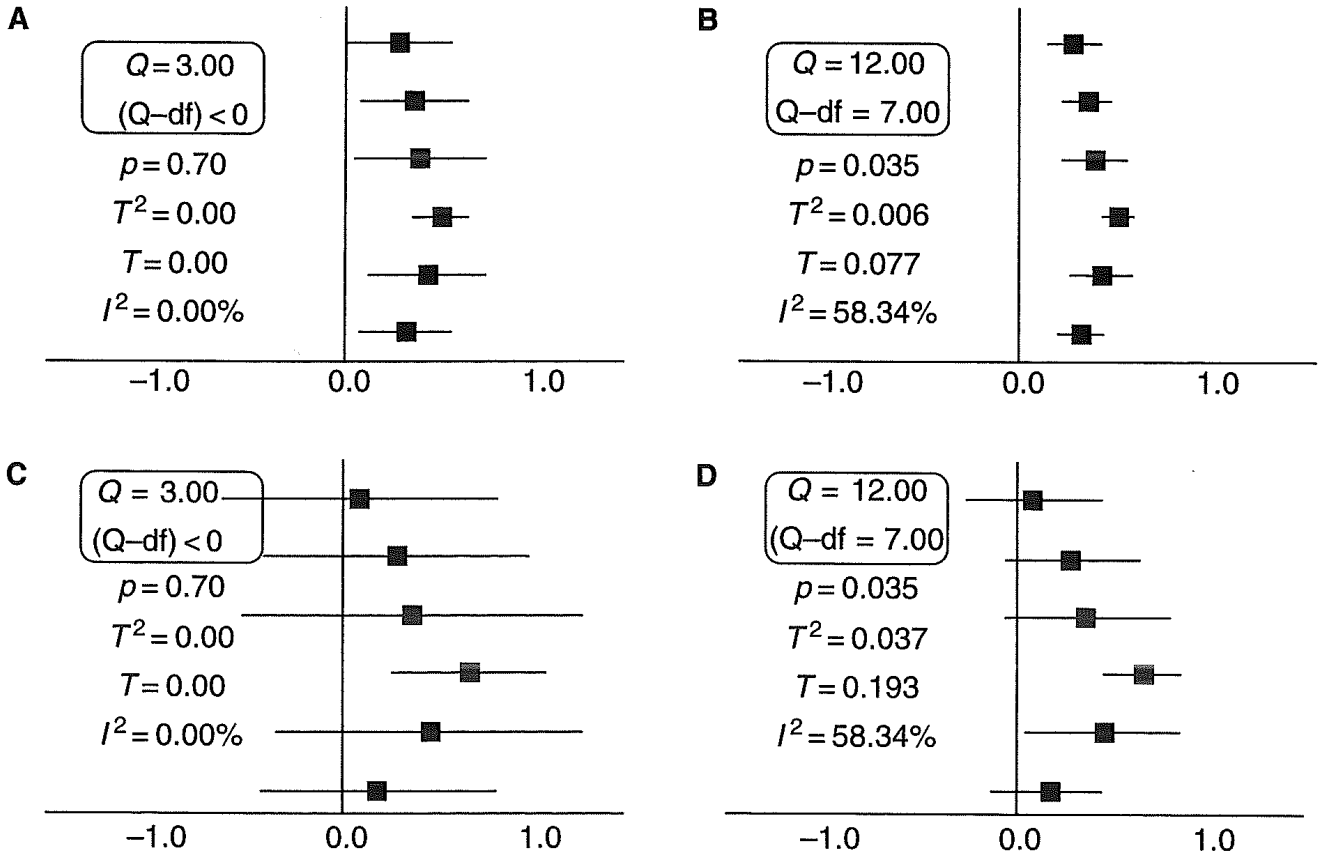
Finnish Institute of Occupational Health

Statistical heterogeneity



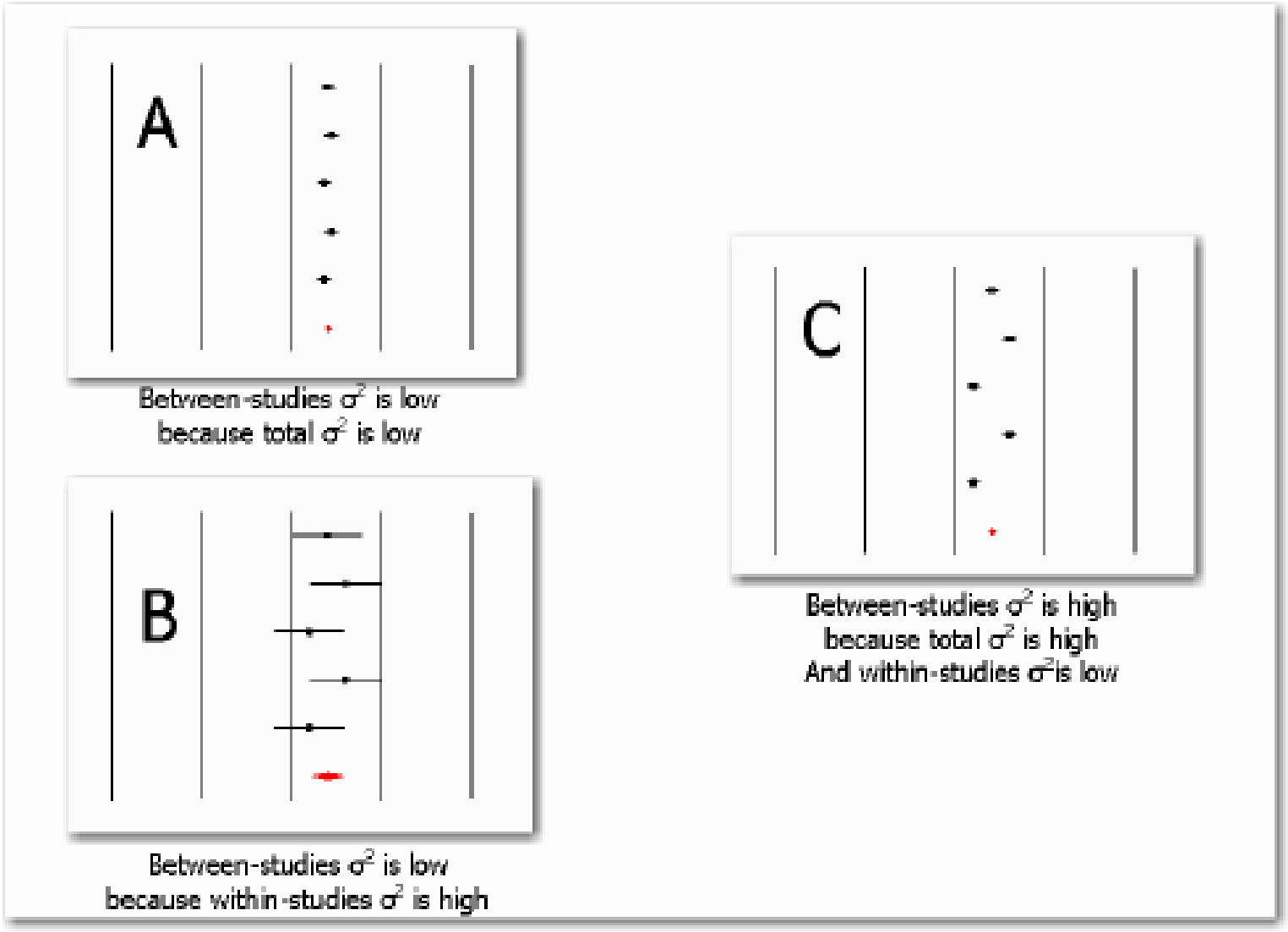
Finnish Institute of Occupational Health

Statistical heterogeneity



Finnish Institute of Occupational Health

Statistical heterogeneity

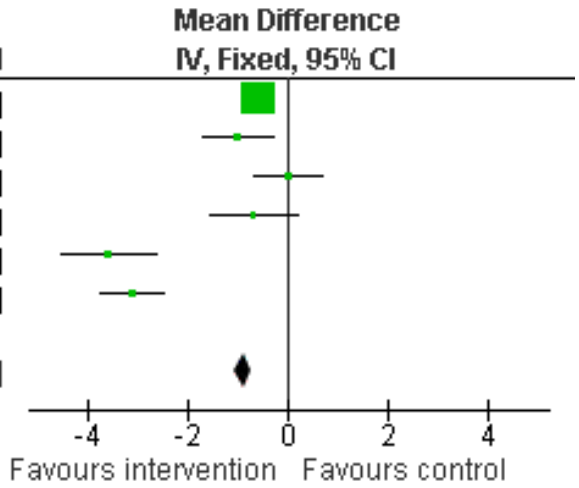


Zinc for Common Cold

Study or Subgroup	Intervention Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Fixed, 95% CI
Kurugol 2006a	4.7	0.8	97	5.3	0.7	97	71.7%	-0.60 [-0.81, -0.39]
Kurugol 2007	5.5	1.97	60	6.5	1.97	60	6.5%	-1.00 [-1.70, -0.30]
Macknin 1998	8.5	2.85	125	8.5	2.85	124	6.4%	0.00 [-0.71, 0.71]
Petrus 1998	4.4	1.4	52	5.1	2.8	49	4.2%	-0.70 [-1.57, 0.17]
Prasad 2000	4.5	1.6	25	8.1	1.8	23	3.4%	-3.60 [-4.57, -2.63]
Prasad 2008	4	1.04	25	7.12	1.26	25	7.8%	-3.12 [-3.76, -2.48]
Total (95% CI)			**384**			**378**	**100.0%**	**-0.89 [-1.07, -0.71]**

Heterogeneity: Chi² = 90.33, df = 5 (P < 0.00001); I² = 94%

Test for overall effect: Z = 9.76 (P < 0.00001)



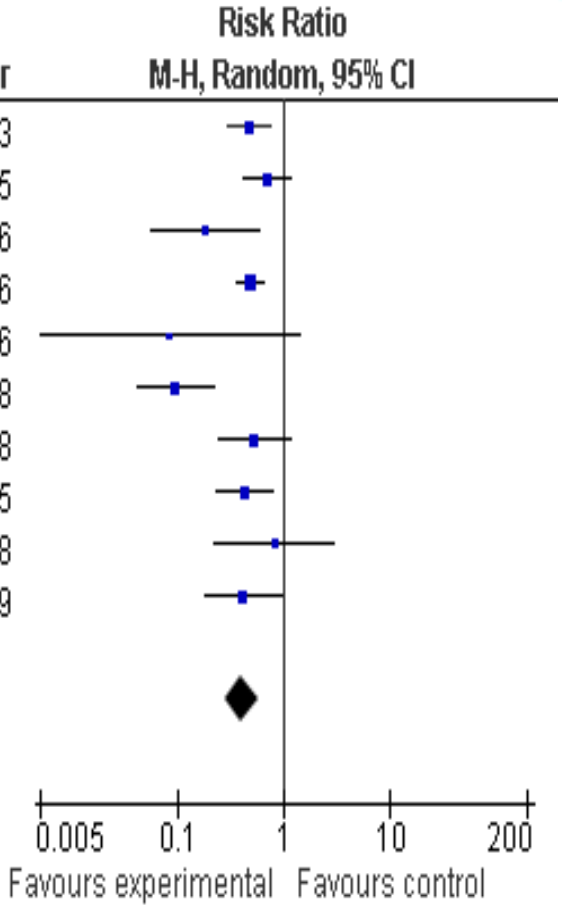
Finnish Institute of Occupational Health

Statistical heterogeneity

Study or Subgroup	Blunted needles Events	Total	Sharp needles Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Year
Wright 1993	18	76	31	62	14.4%	0.47 [0.29, 0.76]	1993
Thomas 1995	14	40	20	40	13.7%	0.70 [0.41, 1.18]	1995
Hartley 1996	3	46	14	39	6.3%	0.18 [0.06, 0.59]	1996
Mingoli 1996	49	390	102	392	16.9%	0.48 [0.35, 0.66]	1996
Rice 1996	0	36	10	64	1.5%	0.08 [0.01, 1.39]	1996
Botet 1998	6	200	63	200	9.7%	0.10 [0.04, 0.21]	1998
Ablett 1998	9	104	15	91	10.2%	0.53 [0.24, 1.14]	1998
Nordkam 2005	12	100	28	100	12.3%	0.43 [0.23, 0.79]	2005
Wilson 2008	4	217	5	221	5.5%	0.81 [0.22, 2.99]	2008
Sullivan 2009	7	97	17	97	9.5%	0.41 [0.18, 0.95]	2009
Total (95% CI)		**1306**		**1306**	**100.0%**	**0.40 [0.28, 0.57]**	
Total events	122		305				

Heterogeneity: $Tau^2 = 0.18$; $Chi^2 = 23.42$, df = 9 (P = 0.005); $I^2 = 62\%$

Test for overall effect: Z = 5.00 (P < 0.00001)



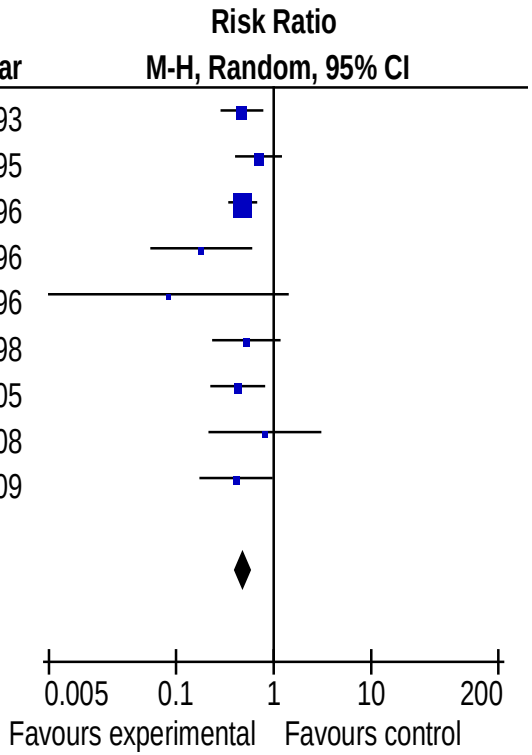
Finnish Institute of Occupational Health

Statistical heterogeneity

Study or Subgroup	Blunted needles Events	Total	Sharp needles Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Year
Wright 1993	18	76	31	62	17.4%	0.47 [0.29, 0.76]	1993
Thomas 1995	14	40	20	40	14.2%	0.70 [0.41, 1.18]	1995
Mingoli 1996	49	390	102	392	40.5%	0.48 [0.35, 0.66]	1996
Hartley 1996	3	46	14	39	2.8%	0.18 [0.06, 0.59]	1996
Rice 1996	0	36	10	64	0.5%	0.08 [0.01, 1.39]	1996
Ablett 1998	9	104	15	91	6.5%	0.53 [0.24, 1.14]	1998
Nordkam 2005	12	100	28	100	10.3%	0.43 [0.23, 0.79]	2005
Wilson 2008	4	217	5	221	2.3%	0.81 [0.22, 2.99]	2008
Sullivan 2009	7	97	17	97	5.6%	0.41 [0.18, 0.95]	2009
Total (95% CI)		**1106**		**1106**	**100.0%**	**0.49 [0.40, 0.59]**	
Total events	116		242				

Heterogeneity: $Tau^2 = 0.00$; $Chi^2 = 7.24$, df = 8 (P = 0.51); $I^2 = 0\%$

Test for overall effect: Z = 7.13 (P < 0.00001)



Finnish Institute of Occupational Health

Statistical heterogeneity

THE COCHRANE COLLABORATION®

Finnish Institute of Occupational Health

Meta-regression

- Linear regression model
  - Dependent variable:
    - effect size (SMD, ln OR)
  - Independent variables
    - any study characteristic (‘subgroup’)
- Tests for differences between ‘subgroups’
- Needs at least 10 studies
- Can be best performed in STATA

Finnish Institute of Occupational Health

Flowchart

List of included studies

Check the conceptual similarity of the items 1 to 7 and in that order

1. Interventions / Exposure
2. Control condition
3. Participants
4. Study Design
5. Outcome
6. Follow-up time
7. Effect Size (RR, MD)

Transform ES if necessary

1. If a little dissimilar, consider making subgroups
2. If quite dissimilar consider narrative synthesis
3. If very dissimilar consider describing studies separately
4. Always report and pool different study designs separately

Perform Meta-Analysis

Check / Explain Remaining Statistical Heterogeneity

Finnish Institute of Occupational Health