Observational studies vs. randomised controlled trials

Mical Paul
Rambam Health Care Campus
Haifa, Israel

ESCMID conference on revival of old antibiotics

Vienna, 22-23 October 2014

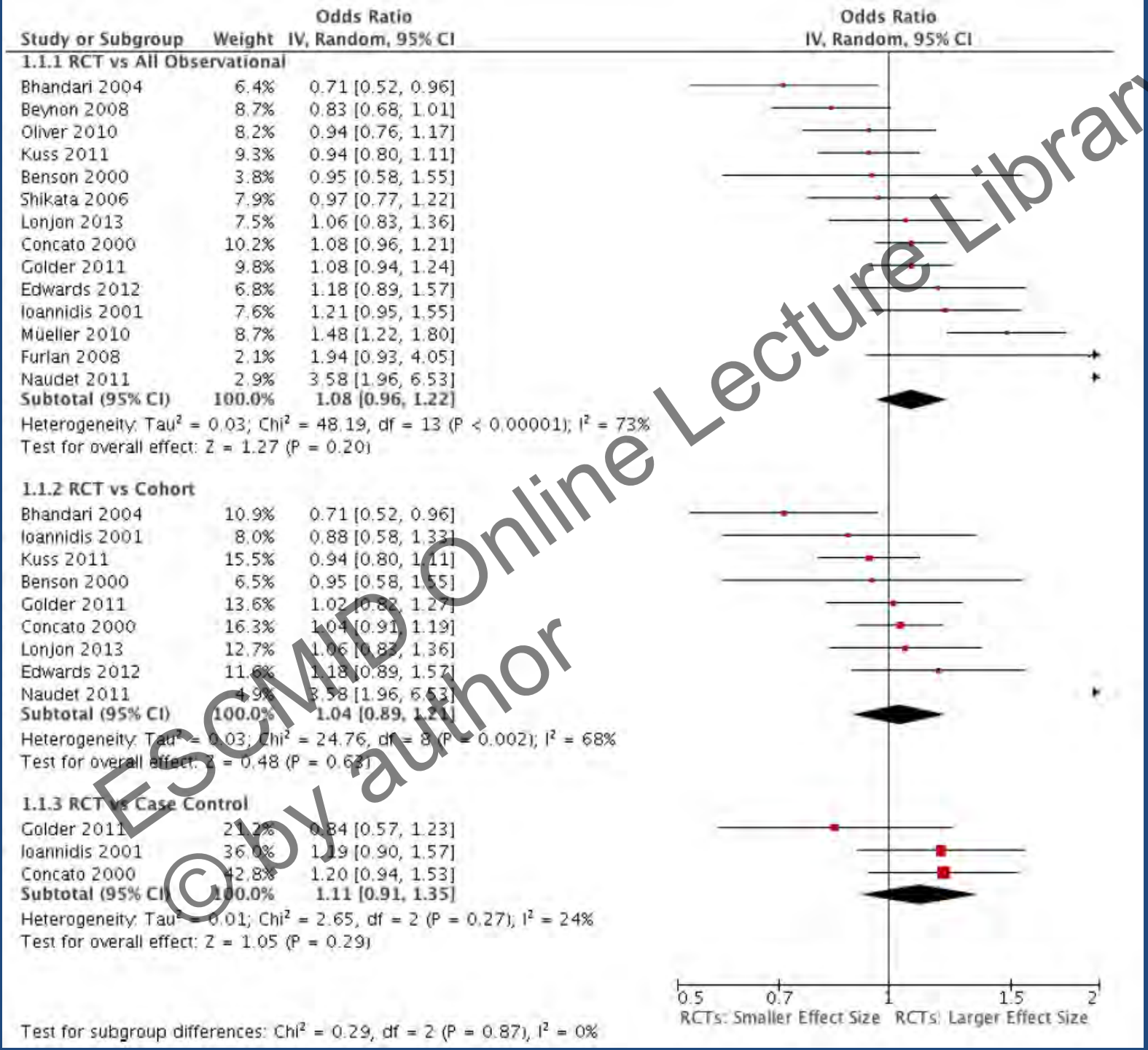
Definitions

- Observational study: Non-interventional study where comparisons or associations are based on observed data
  - Treatment was prescribed based on clinical considerations
- Randomized controlled trial: Treatment assignment is based on a random code such that the treatment assignment of the next patient cannot be predicted

What are we asking

- Is bias in observational studies adjustable so as to approximate treatment effects of RCTs?

- Does poor external validity of RCTs affect their effect estimates



Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials

Anglemyer et al. Cochrane Database of Systematic reviews 2014

Characteristics of methodological studies

- Matched observational studies to RCTs comparing same/ similar interventions
- Compared effect estimates of RCTs to adjusted effect estimates from the observational studies for the same outcome
- Studies differed on inclusion criteria, matching of patient characteristics, and whether individual comparisons were reported separately or combined

Chalmers et al. N Engl J Med 1983

- Restriction on study inclusion: compared blinded RCTs vs. non-blinded RCTs vs. observational
- Matched patient characteristics: no, but examined distribution of confounders
- Comparisons: only treatments for acute MI included
- Absolute difference between intervention and control of 8.8% in blinded RCTs, 24.4% of open-label RCTs, and 58.1% of the observational studies
- At least one prognostic variable was maldistributed ($p<0.05$) in 14%, 26.7% and 58.1% of studies, respectively

Sacks et al. Am J Med 1985

- Restriction on study inclusion: only observational studies using historical controls were included
- Matched patient characteristics: no
- Comparisons: pooled together
- The agent being tested was considered more effective than control in 44 of 56 trials with historical controls (79%), but in only 10 of 50 RCTs (20%)
- Outcomes rates were similar with the intervention in RCTs and observational studies, but significantly more frequently in the control arm in observational studies

Benson et al. N Engl J Med 2000

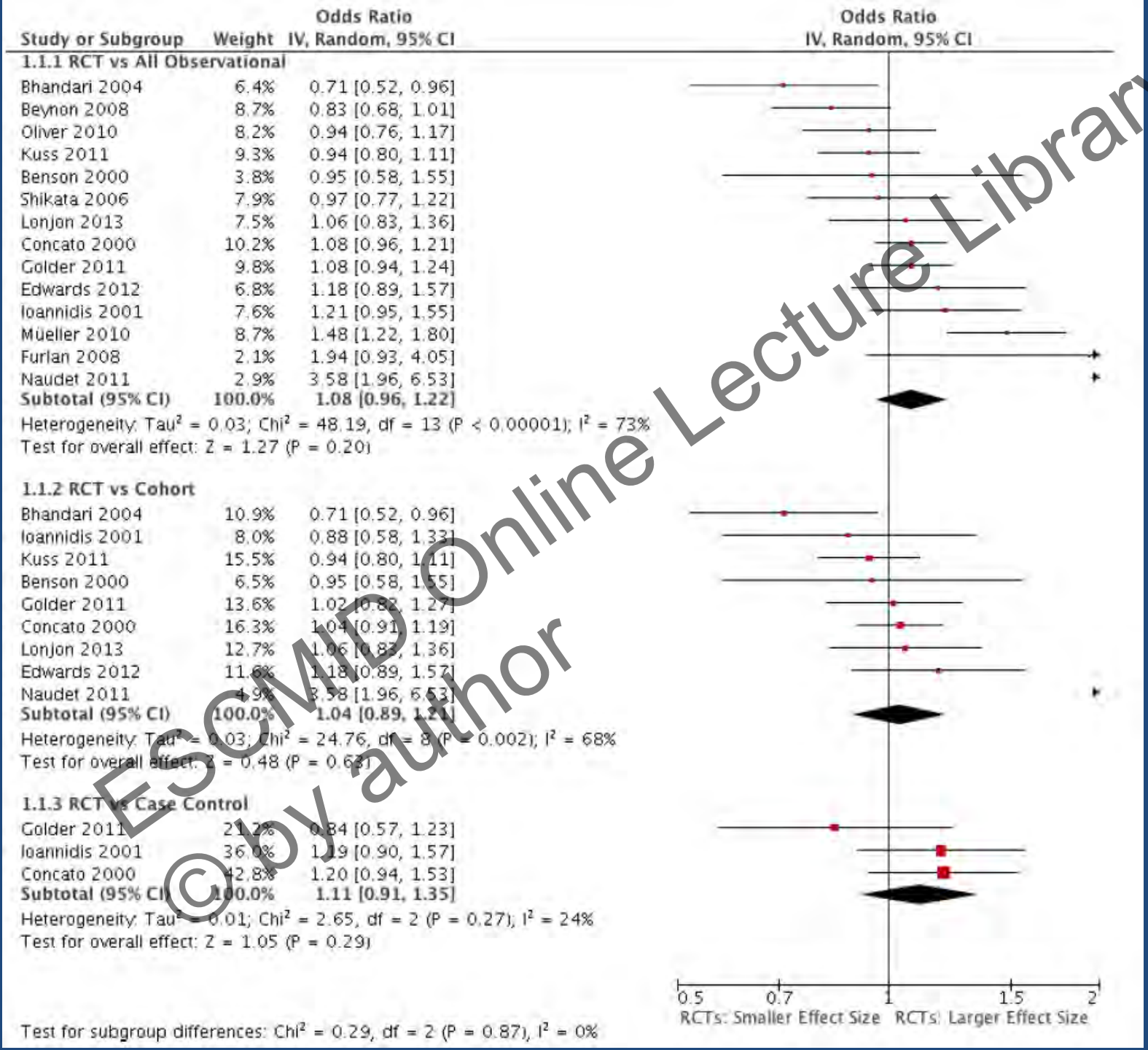
- Restriction on study inclusion: published in the 120 most widely read journals
- Matched patient characteristics: partially by matching inclusion criteria
- Comparisons: pooled separately
- Out of 19 comparisons, in 2 a statistically significant difference between RCTs and observational studies was observed
  - In 1 effects larger in a observational study: PTCA better than CABG
  - In1 effect larger in a RCT: treatment of retinal detachment

Concato et al. N Engl J Med 2000

- Restriction on study inclusion: five major medical journals
- Matched patient characteristics: no
- Comparisons: pooled separately
- Five comparisons were assessed and results were similar in RCTs and observational studies
- Confidence in results was smaller in RCTs

Ioannidis et al. JAMA 2001

- Restriction on study inclusion: all studies were included in meta-analysis
- Matched patient characteristics: no
- Comparisons: pooled separately
- Out of 45 comparisons, in 7 a statistically significant difference between RCTs and observational studies was observed
  - In 6 effects larger in observational studies: CABG vs. medical Tx for CAD, anticoagulants in MI, screening mammography, Tx for habitual abortions or alcohol dependence, and prevention of STDs
  - In1 effect larger in a RCT comparing two surgical interventions for urinary incontinence
- Within the same comparison, observational studies' results were more heterogeneous than RCT results



Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials

Anglemyer et al. Cochrane Database of Systematic reviews 2014

Naudet et al. PLoS One. 2011

- Restriction on study inclusion: none, published and unpublished studies included
- Matched patient characteristics: yes, meta-regression
- Comparisons: only fluoxetine or venlafaxine in the first-line treatment for depressive disorders included
- Intervention effects, measured as the difference between pre and post depression scores, were greater by a magnitude of 4.59 (2.61 to 6.56) in RCTs compared to the observational studies

Bhandari et al. Arch Orthop Trauma Surg. 2004

- Restriction on study inclusion: none
- Matched patient characteristics: no
- Comparisons: only internal fixation vs. arthroplasty in patients with femoral neck fracture included
- Observational studies yielded larger effects for mortality and the need for revision by 40% and 19%, respectively

Papanikolaou 2006 CMAJ

- Restriction on study inclusion: only studies with >4000 patient included
- Matched patient characteristics: no
- Comparisons: all pooled. Only effects of interventions on adverse effects assessed
- Larger effects for harms in RCTs
- The estimated increase in RR differed by more than two-fold between observational studies and RCTs for 54% of the topics studied

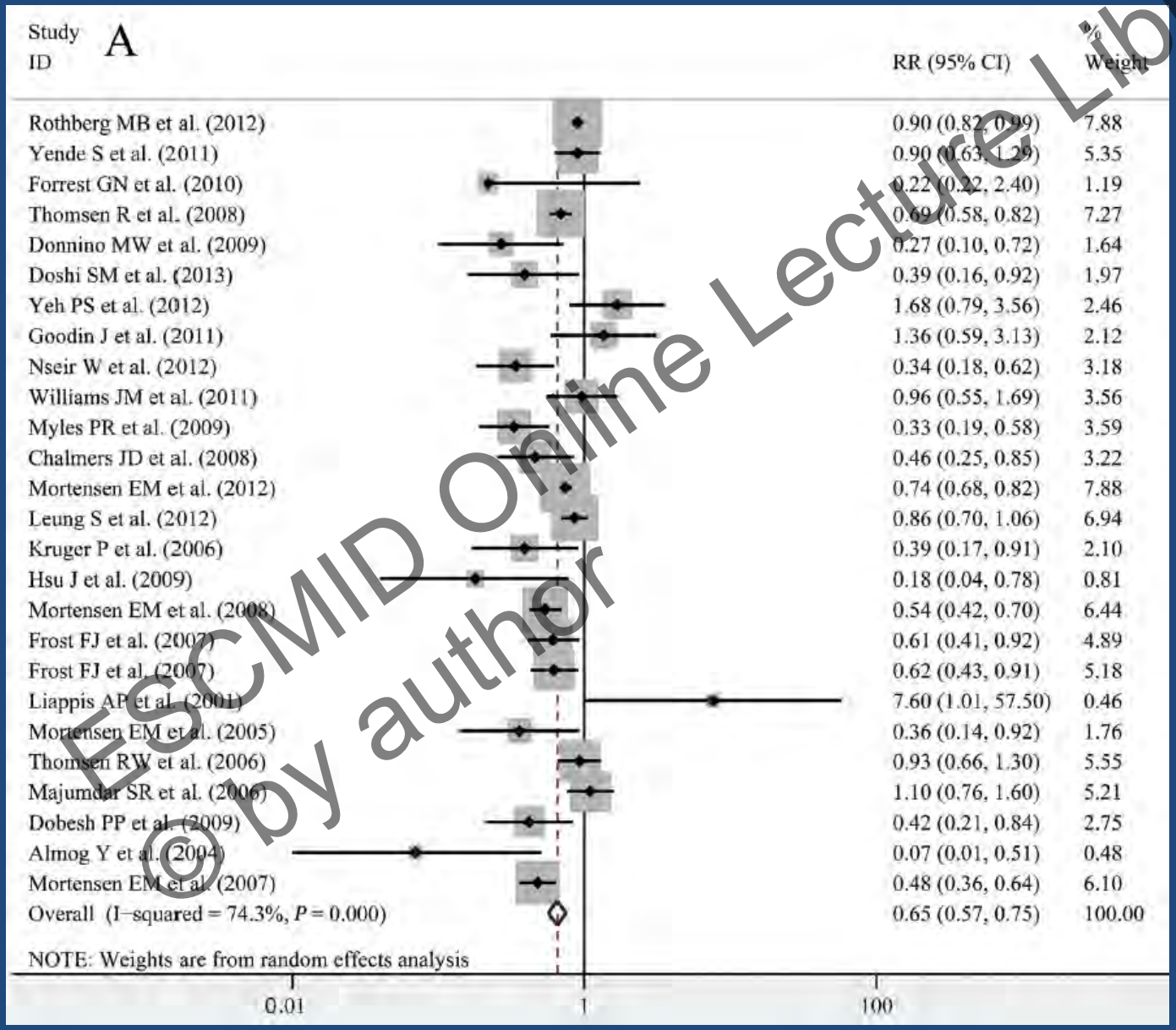
Sources of heterogeneity Comparing observational to RCTs

- The likelihood of selection bias in the question of interest
  - Methods and ability of adjustment
- The outcome selected and the “interest” of showing an advantage or disadvantage
- Differences between patient characteristics between the RCTs and the observational studies

What about RCT vs. observational studies in infectious diseases?

None of the reviews examined separately interventions in infectious diseases/ antibiotics

Statins for sepsis – observational studies



Adjusted effect estimates

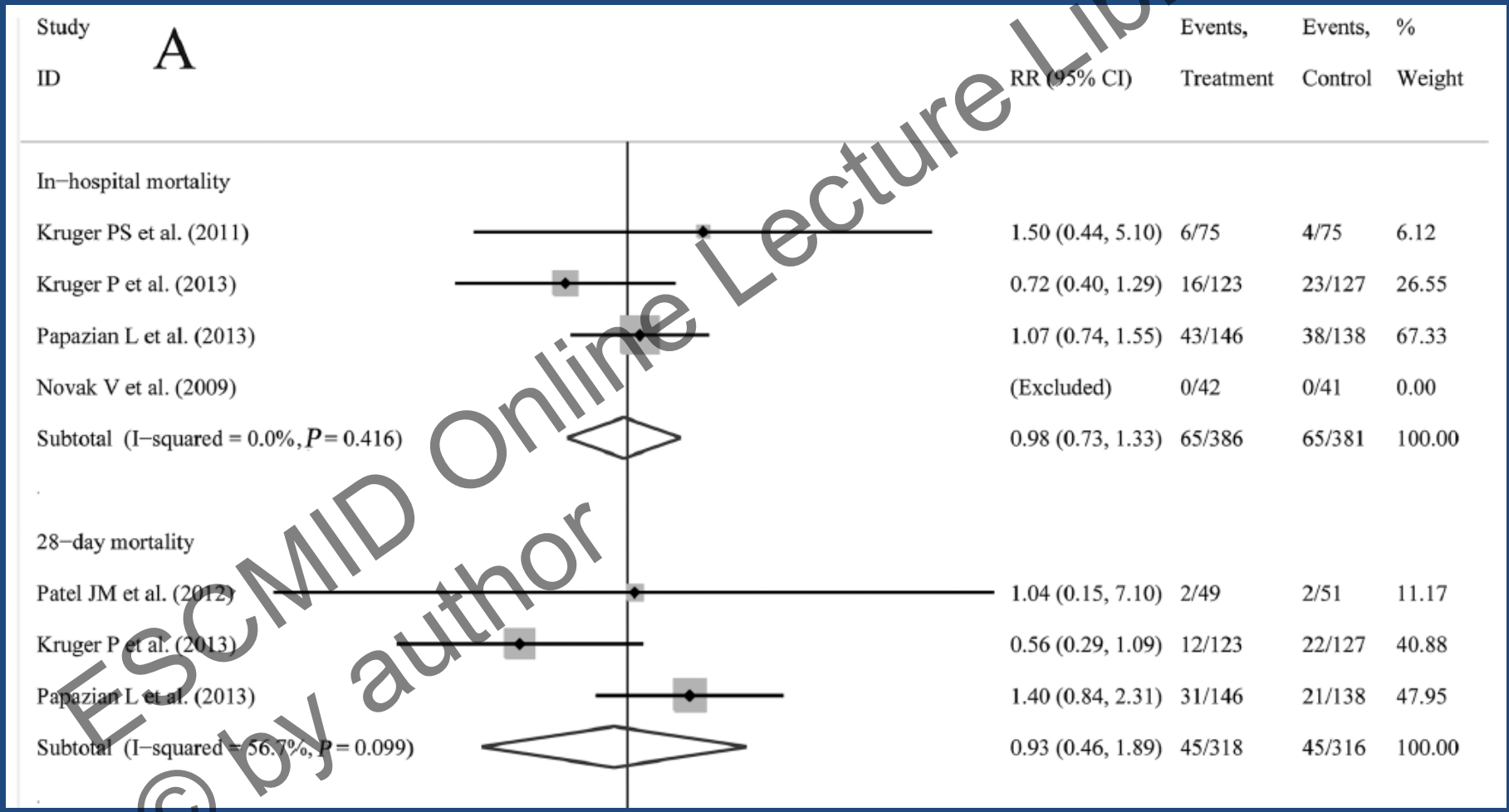
Pooled RR 0.65 (0.57-0.75)

Statins for sepsis – RCTs

?

Wan et al. Crit Care. 2014

Statins for sepsis – RCTs



Pooled RRs 0.98 (0.73-1.33) and 0.92 (0.46-1.89)

Problem ?

Mostly bias: healthy user bias

Neuraminidase inhibitors for influenza: observational studies

Table 2: NAI TREATMENT (AT ANY TIME) VS. NONE

Subgroups	Crude analysis		Adjusted† analysis	
	OR (95% CI)	P value	OR (95% CI)	P value
Laboratory confirmed or clinically diagnosed (all ages); n=29,234	0·92 (0·81 to 1·05)	0·21	0·81 (0·70 to 0·93)	0·0024
Laboratory confirmed cases (all ages) ; n=25,001	0·94 (0·81 to 1·09)	0·42	0·82 (0·70 to 0·95)	0·0104
Adults (16 years and above) ; n=19,816	0·82 (0·70 to 0·95)	0·0071	0·75 (0·64 to 0·87)	0·0002
Children (below 16 years); n=9,218	1·02 (0·73 to 1·42)	0·90	0·82 (0·58 to 1·17)	0·28
Pregnant women; n=2,166	0·47 (0·24 to 0·90)	0·0228	0·46 (0·23 to 0·89)	0·0215
ICU patients				
Adults (≥16 years); n=5,103	0·74 (0·57 to 0·95)	0·0187	0·72 (0·56 to 0·94)	0·0155
Children (<16 years); n=1,725	0·84 (0·52 to 1·37)	0·49	0·70 (0·42 to 1·16)	0·17

†adjusted for treatment propensity (by quintile), corticosteroid use and antibiotic use

Muthuri et al. Lancet Resp Med 2014

Neuraminidase inhibitors for influenza: observational studies

Subgroups	Crude analysis		Adjusted† analysis	
	OR (95% CI)	P value	OR (95% CI)	P value
Early treatment vs. none:				
Laboratory confirmed or clinically diagnosed (all ages); n=16,425	0·54 (0·40 to 0·72)	<0·0001	0·50 (0·37 to 0·67)	<0·0001
Laboratory confirmed cases (all ages); n= 13,200	0·53 (0·39 to 0·71)	<0·0001	0·48 (0·36 to 0·66)	<0·0001
Adults (16 years and above); n=10,607	0·39 (0·28 to 0·55)	<0·0001	0·38 (0·27 to 0·54)	<0·0001
Children (below 16 years); n=5,696	1·08 (0·61 to 1·93)	0·79	0·85 (0·47 to 1·53)	0·59
Pregnant women, n=1,303	0·16 (0·04 to 0·64)	0·0099	0·16 (0·04 to 0·67)	0·0118
ICU patients				
Adults (≥16 years); n=1,608	0·30 (0·19 to 0·45)	<0·001	0·31 (0·20 to 0·47)	<0·001
Children (≤16 years); n=572	0·88 (0·40 to 1·91)	0·74	0·76 (0·34 to 1·67)	0·49

†adjusted for treatment propensity (by quintile), corticosteroid use and antibiotic use

Neuraminidase inhibitors for influenza: observational studies

Early treatment vs. Later treatment:

Subgroups	Crude analysis		Adjusted† analysis	
	OR (95% CI)	P value	OR (95% CI)	P value
Laboratory confirmed or clinically diagnosed (all ages); n=13,254	0·36 (0·31 to 0·41)	<0·0001	0·48 (0·41 to 0·56)	<0·0001
Laboratory confirmed cases (all ages) ; n=12,992	0·36 (0·31 to 0·41)	<0·0001	0·48 (0·41 to 0·56)	<0·0001
Adults (16 years and above); n=9,270	0·37 (0·32 to 0·44)	<0·0001	0·45 (0·38 to 0·54)	<0·0001
Children (below 16 years); n=3,899	0·53 (0·35 to 0·80)	0·0026	0·67 (0·44 to 1·03)	0·07
Pregnant women ; n= 917	0·20 (0·09 to 0·46)	0·0002	0·27 (0·11 to 0·63)	0·0026
ICU patients				
Adults (≥16 years); n=3,385	0·64 (0·51 – 0·79)	<0·0001	0·62 (0·49 to 0·77)	<0·0001
Children (<16 years); n=683	1·12 (0·63 to 1·99)	0·69	1·15 (0·64 to 2·06)	0·64

Neuraminidase inhibitors for influenza: RCTs

?

Neuraminidase inhibitors for influenza: RCTs

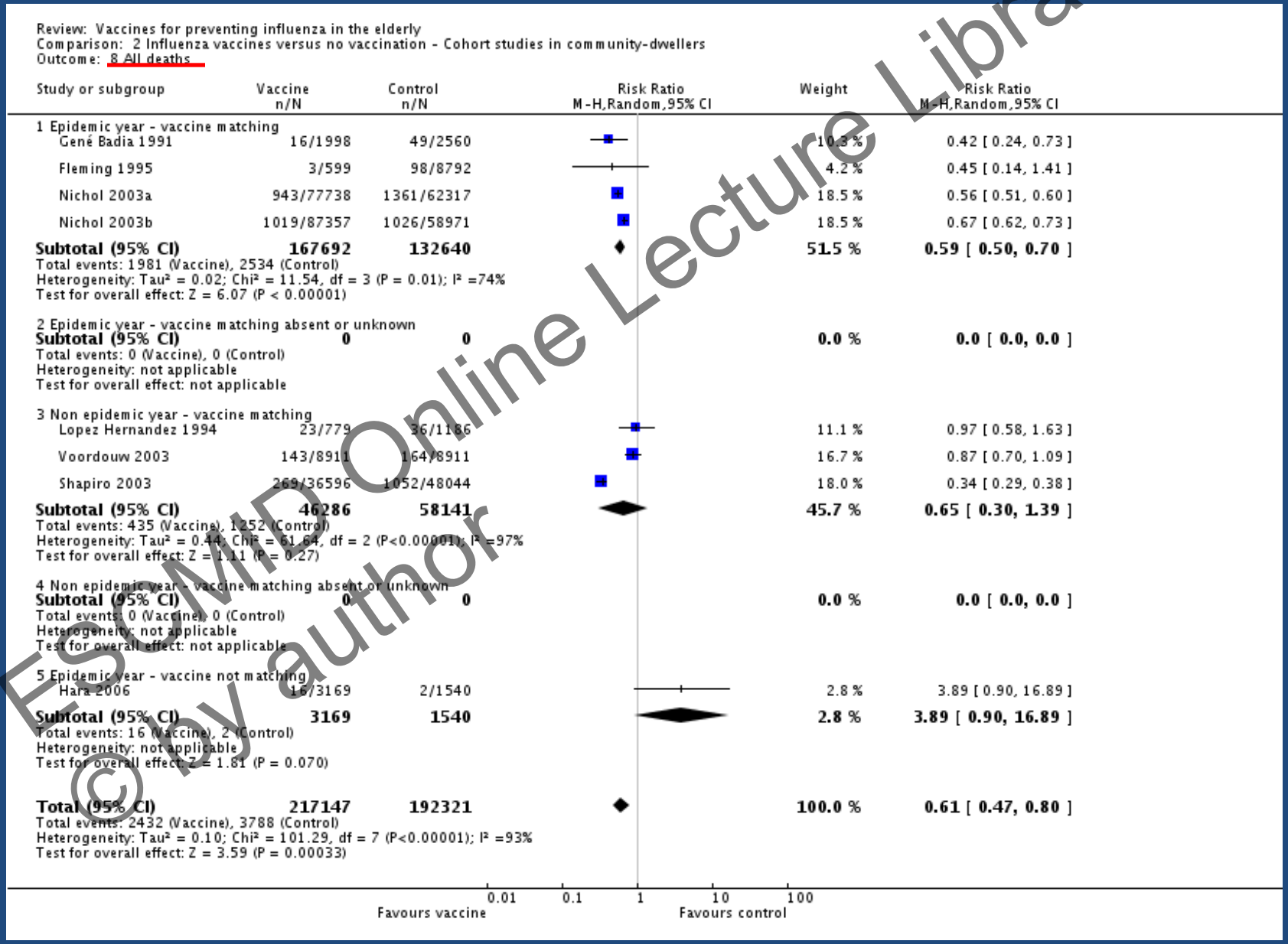
- 49 RCTs included in the meta-analysis
- Oseltamivir reduced the duration of symptoms by 16.8 hours (8.4-25.1)
- Hospital admissions were identical in the placebo and NAI treatment arms
- Oseltamivir significantly reduced self reported, investigator-mediated, unverified pneumonia by 1.00% (0.22 to 1.49%)
  - The effect was not significant in the five trials that used a more detailed diagnostic form for pneumoniaNausea and vomiting were the more common with oseltamivir
- Oseltamivir significantly increased nausea and vomiting

Jefferson et al. Cochrane Database Syst Rev. 2014

Problem ?

Mostly external validity: different patient groups. Patients in the community vs. hospitalized patients

Influenza vaccines: observational studies



Influenza vaccines: observational studies

- Vaccines also associated with lower rates of ILIs, pneumonia, hospitalizations (mainly in nursing home patients and deaths from influenza

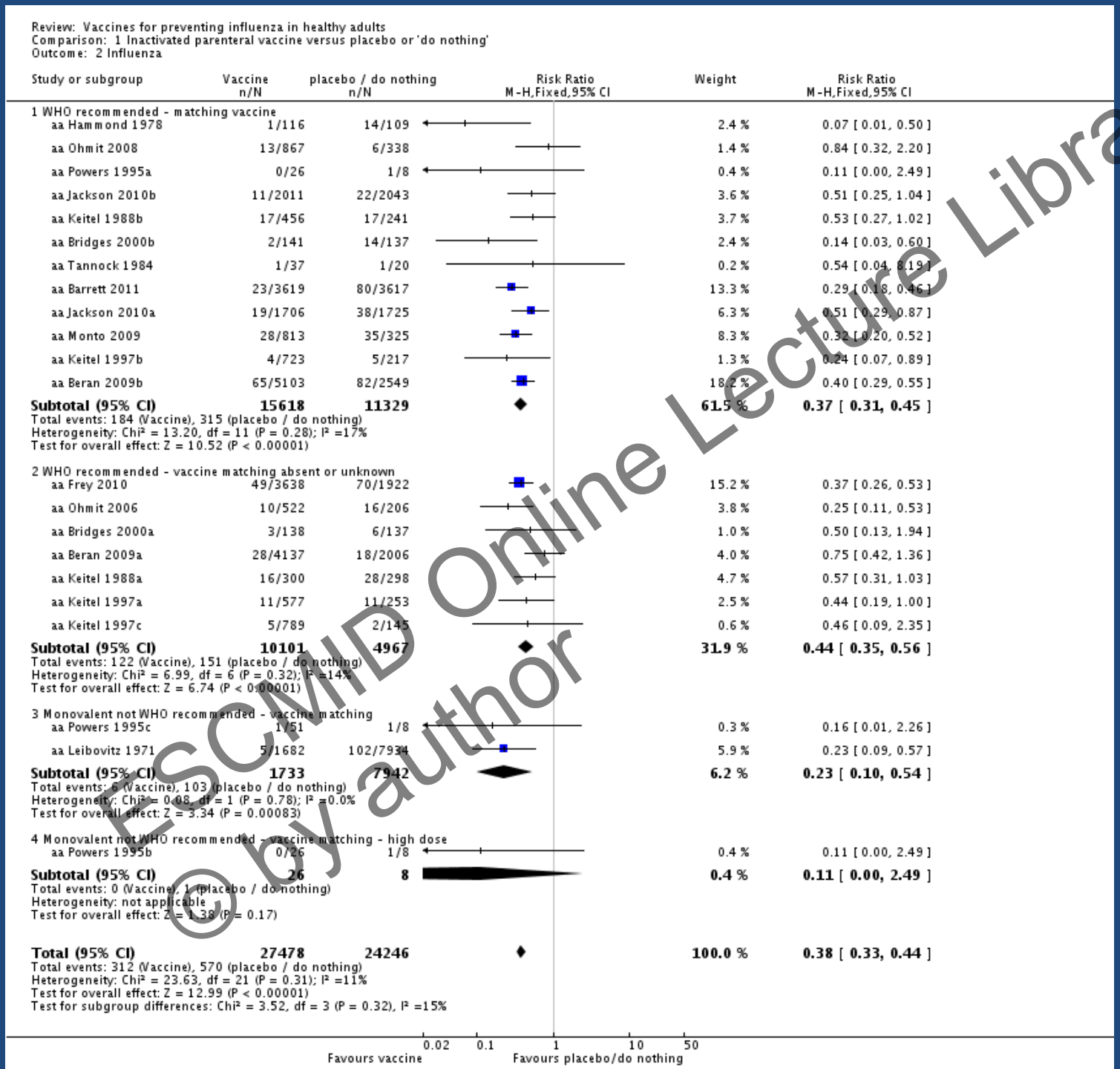
Jefferson et al. Cochrane Database Sys Rev 2010

Influenza vaccines: RCTs

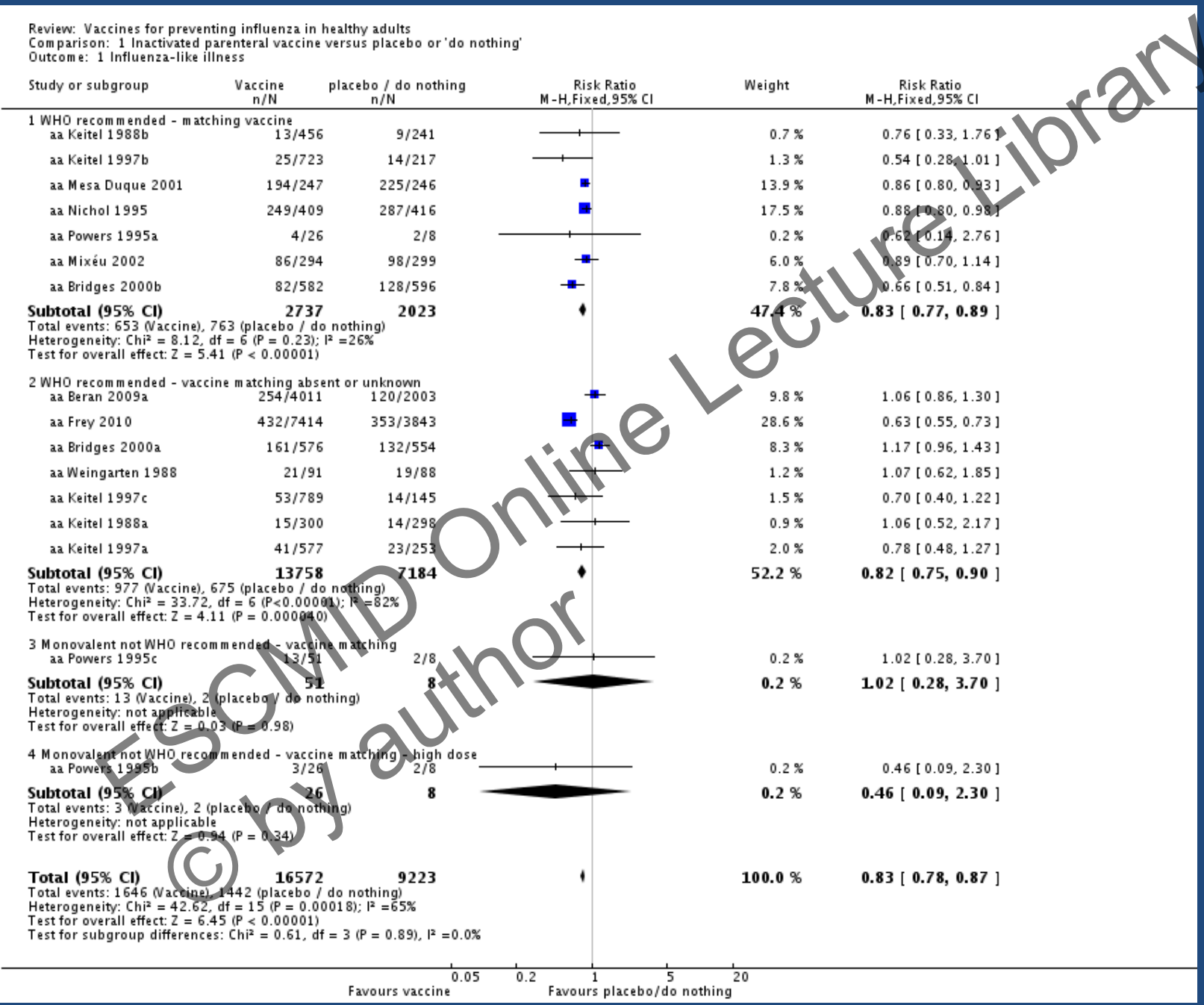
?

Influenza

Influenza vaccines: RCTs



Review: Vaccines for preventing influenza in healthy adults
Comparison: 1 Inactivated parenteral vaccine versus placebo or 'do nothing'
Outcome: 1 Influenza-like illness



Influenza-like illness

Influenza vaccines: RCTs

Influenza vaccines: RCTs

- Mortality not reported
- modest effect on time off work
- No effect on hospital admissions or complication rates

Jefferson et al. Cochrane Database Sys Rev 2014

Problem ?

Mostly external validity: different patient groups. Healthy adults vs. elderly people

Cotrimoxazole vs. vancomycin for invasive MRSA infections: patient characteristics

	RCT included N=252	RCT excluded N=220	P value
Urine catheter prior to infection	80 (31.7%)	138 (62.7%)	<0.001
Central venous catheter prior to infection	32 (12.7%)	104 (47.2%)	<0.001
Mechanical ventilation at onset	27 (10.7%)	98 (44.5%)	<0.001
Platelets, K/m^3 (median, percentile)	279 (192-403)	218 (123-331)	<0.001
Urea, mg/dl (median, percentile)	38 (25-63)	55 (34-101)	<0.001
Albumin, g/dl (median, percentile)	2.8 (2.4-3.3)	2.7 (2.1-3.3)	0.059
Septic Shock at onset	6 (2.4%)	23 (10.5%)	<0.001
SOFA score at onset (median, percentile)	2 (1-4)	3 (2-4)	<0.001

Problem ?

External validity and bias: different patient groups. Patients included vs. excluded (no informed consent); difficult to adjust for "healthy" sick patients given TMP-SMX

Macrolides for CAP

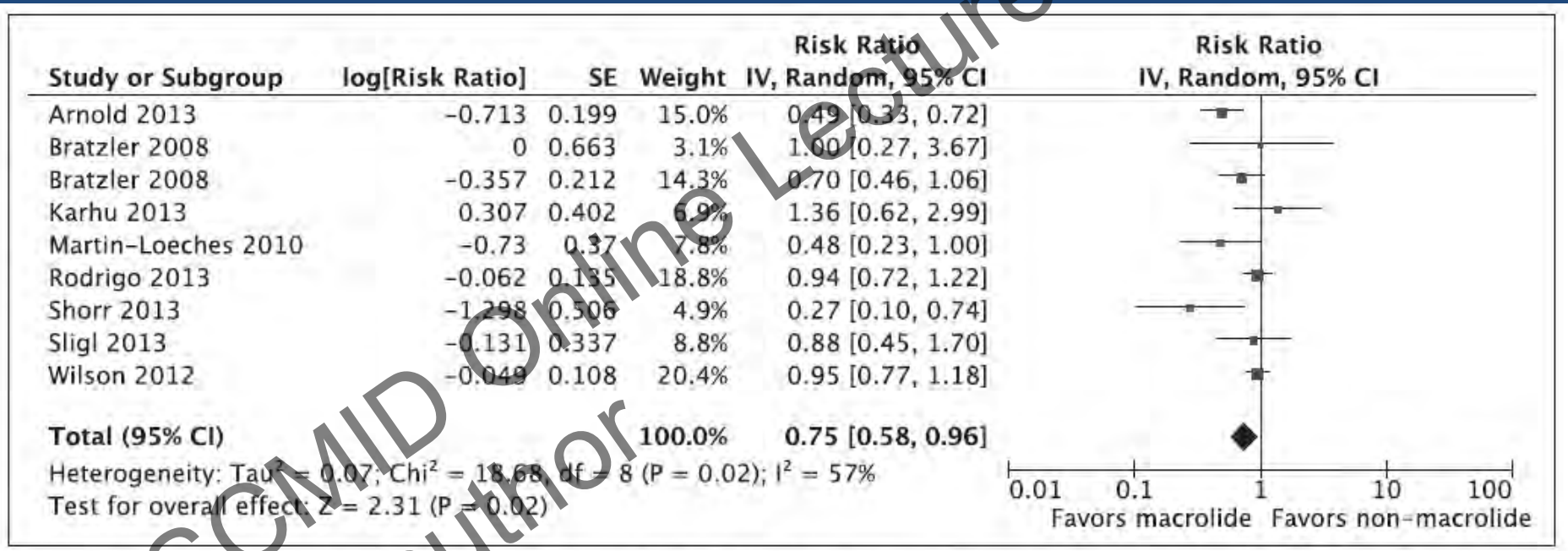


Figure 3. Macrolide versus nonmacrolide therapy and mortality in critically ill patients with community-acquired pneumonia: pooled adjusted risk estimates ($n = 9$).

Sligl et al. Crit Care Med. 2014

Macrolides for CAP: RCTs

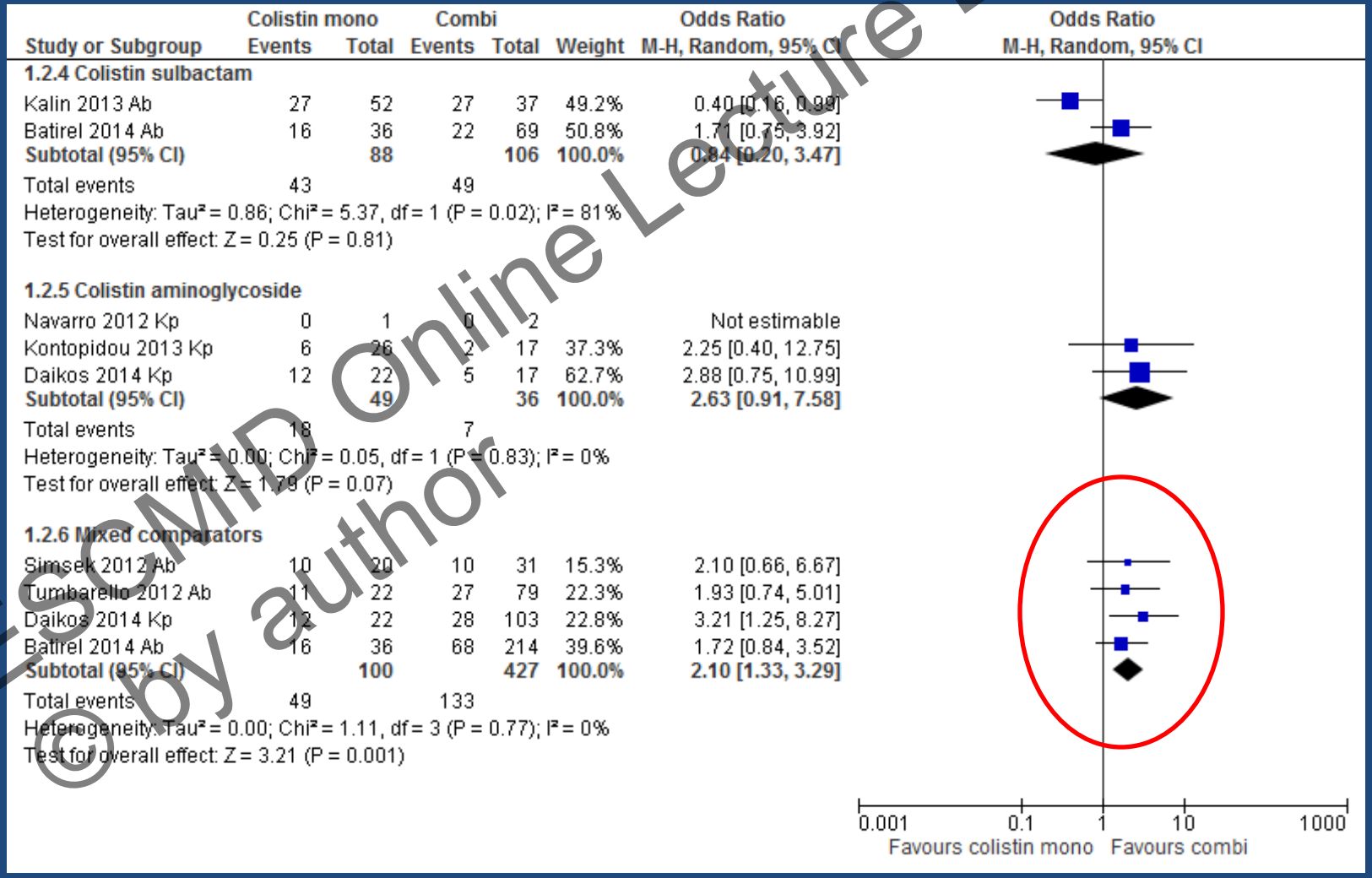
?

Antibiotic treatment of moderate-severe community-acquired pneumonia: design and rationale of a multicentre cluster-randomised cross-over trial.

- Multi-centre cluster randomised cross-over study, comparing empirical antibiotic strategies.
- Participating centres are randomised to three consecutive periods of four months, in which one of the three empirical antibiotic strategies applies.
- In each hospital the local antibiotics committee has been asked to adopt this empirical strategy as the standard treatment for CAP during that period.
- Consequently, written informed consent is not needed prior to the start of the preferred treatment of the study, but only for collection of individual patient data.

van Werkhoven CH1, Postma DF, Oosterheert JJ, Bonten MJ. Neth J Med. 2014 Apr

Colistin-carbapenem combination therapy: observational studies



Study or Subgroup	Colistin mono Events	Total	Combi Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
1.2.4 Colistin sulbactam						
Kalin 2013 Ab	27	52	27	37	49.2%	0.40 [0.16, 0.98]
Batirel 2014 Ab	16	36	22	69	50.8%	1.71 [0.75, 3.92]
Subtotal (95% CI)		**88**		**106**	**100.0%**	**0.84 [0.20, 3.47]**
Total events	43		49			
Heterogeneity: Tau² = 0.86; Chi² = 5.37, df = 1 (P = 0.02); I² = 81%						
Test for overall effect: Z = 0.25 (P = 0.81)						
1.2.5 Colistin aminoglycoside						
Navarro 2012 Kp	0	1	0	2		Not estimable
Kontopidou 2013 Kp	6	26	2	17	37.3%	2.25 [0.40, 12.75]
Daikos 2014 Kp	12	22	5	17	62.7%	2.88 [0.75, 10.99]
Subtotal (95% CI)		**49**		**36**	**100.0%**	**2.63 [0.91, 7.58]**
Total events	18		7			
Heterogeneity: Tau² = 0.00; Chi² = 0.05, df = 1 (P = 0.83); I² = 0%						
Test for overall effect: Z = 1.79 (P = 0.07)						
1.2.6 Mixed comparators						
Simsek 2012 Ab	10	20	10	31	15.3%	2.10 [0.66, 6.67]
Tumbarello 2012 Ab	11	22	27	79	22.3%	1.93 [0.74, 5.01]
Daikos 2014 Kp	12	22	28	103	22.8%	3.21 [1.25, 8.27]
Batirel 2014 Ab	16	36	68	214	39.6%	1.72 [0.84, 3.52]
Subtotal (95% CI)		**100**		**427**	**100.0%**	**2.10 [1.33, 3.29]**
Total events	49		133			
Heterogeneity: Tau² = 0.00; Chi² = 1.11, df = 3 (P = 0.77); I² = 0%						
Test for overall effect: Z = 3.21 (P = 0.001)						

Colistin-carbapenem combination therapy: RCTs

?

Multicenter open-label RCT to compare colistin alone vs. colistin plus meropenem

- Ongoing, in Israel, Greece and Italy
- Targeted recruitment of 360 patients 1:1 randomization
- Primary outcome composite of survival, hemodynamic and clinical stability on day 14
- Among secondary outcomes: 28-day all-cause mortality

Grant Health-F3-2011-278348

The scientific method of Galileo has been subverted before our very eyes.

Galileo observed, described and then produced a hypothesis or theory which he then proceeded to test with an experiment. This model has served us well in the last 400 years with a few exceptions (the already cited penicillin for example).

What we now witness is a fundamental subversion of the order (I am not going to call it a paradigm shift) of things. Observations are fact, case-controls, case series, cohorts (even retrospective and datalinked ones) are being held out as proof. Trials are for regulators, they say.

The origin of all this is complex and partly known. In the Tamiflu story as we began uncovering the extent of reporting bias affecting the clinical trials that had been used to make policy and justify stockpiling, decision makers turned to observational evidence (of universally recognised poor quality) as props for their unchangeable policies.

It is a sad parable of the world we live in.
Best wishes,
Tom.

EVIDENCE-BASED-HEALTH@JISCMAIL.AC.UK
Email discussion lists for the UK Education and Research communities

The definition of breakthrough is "it costs a packet" - and sofosbuvir fits the picture.

This kind of bullshit is replicated in the EU with so called early assessment to get better, innovative drugs earlier to patients who desperately need them. So the burden of proof is slowly being pushed back to phase IV or beyond which may be observational, subverting Galileo's methods.

What we should always remember is that Einstein's general relativity theory (1915) was a theory and remained a theory until Eddington's natural experiment during the 1919 solar eclipse confirmed that gravitation could deflect starlight as the theory had put forward.

The rise of observational data (even non comparative) is an involution, not an evolution. There are many culprits most of them in my profession (I am a physician) and they will be held to account.

Greed and science do not mix.

Nite from Rome.
Tom.

Correspondence Tom Jefferson to RoyrPoses
EVIDENCE-BASED-HEALTH@JISCMAIL.AC.UK

My conclusions

- Only RCTs can provide unbiased treatment comparisons for antibiotics
- The external validity in RCTs of antibiotics is very low, with the need for informed consent being a major contributor to this
- The advantage of old drugs is that they are already used
  - Should be reflected in their testing
  - The informed consent process should be bended

Informed Consent for Comparative Effectiveness Research?

Faden et al. Informed consent, comparative effectiveness, and learning health care. N Engl J Med. 2014 Feb 20;370(8):766-8.

Informed Consent for Comparative Effectiveness Research?

- Testing of commonly used interventions where practice varies
- Patients in the clinical setting are not consulted before treatment selection
- Risks and uncertainty small and interventions do not have significant implications on patients' lives
- Patient participation
  - Helping to set the CER priorities of the system and by serving on ethics-oversight panels that will review proposed CER studies
- Mature learning health care system grounded in a set of moral commitments
- Notification to affected patients, who will have an option to decline participation

Faden et al. Informed consent, comparative effectiveness, and learning health care. N Engl J Med. 2014 Feb 20;370(8):766-8.

Thomas C. Chalmers

- Tom Chalmers (1917-1995) was an outspoken advocate of randomised trials… He is perhaps best known for the notion 'randomise the first patient', his belief that it is more ethical to randomise patients than to treat them in the absence of good evidence.
- *"One of the big myths is that it (RCT) is very expensive. In fact, randomization costs a few cents a patient. What is expensive is observation and recording data. When you go to all the effort to randomize, you make an extra effort to find out what happened to the patients – what other things they received and whether they got their medicine or not. Those are all things you should be doing anyway…"*

Thank you

Conflicts of interests: in the next observational study that we'll publish, we 'll do our best to convince you that observational studies are the best design to test our intervention