COMPARING MULTIPLE INTERVENTIONS WORKSHOP 1

Lorne Becker, Tianjing Li, Chris Schmid

20th Cochrane Colloquium
Auckland, New Zealand
October 1, 2012

Disclosures

- Some of the work is funded by the Cochrane Collaboration Innovation Funds
- None of the presenters have any other conflicts of interest to disclose

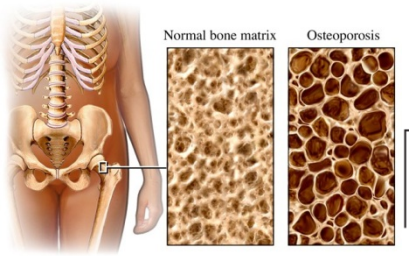
Learning Objectives

- Describe what is an indirect comparison and a network meta-analysis
- Describe utilities of network meta-analysis for comparing multiple interventions
- Gain awareness about current methodologic challenges, statistical complexities, and common errors in the literature when multiple interventions are compared
- Understand the choice of appropriate review type (e.g., intervention review or overview) for the right question

Key Messages

- Network meta-analysis is an extension of standard, pair-wise meta-analysis.
- Use of network meta-analysis is often necessary for drawing inference about multiple competing interventions and a formal approach is preferable.
- The evolution of these methods has led us to re-evaluate the role of Overviews when comparing multiple interventions.

Which Treatment Should be Recommended?



The NEW ENGLAND JOURNAL of MEDICINE

CLINICAL THERAPEUTICS

A 67-year-old woman was referred by her primary care physician for treatment of osteoporosis and progressive bone loss. One year before the visit, the patient had discontinued hormone-replacement therapy. She had subsequently begun to experience midback pain and lost 1.5 inch in height. A x-ray scan has confirmed a diagnosis of osteoporosis. One year later, a second scan showed a further decrease of bone mineral density at the lumbar spine, as well as a compression fracture of the 11th thoracic vertebra.

Which treatment should be recommended?

Paraphrased from
Favus NEJM 2010

Treatment of Osteoporosis and Risk of Hip Fracture

- **Medical treatment:**

  Over 10 drugs/combination of drugs

  - ✓ Estrogen
  - ✓ Selective estrogen receptor modulators (SERMs)- Raloxifene
  - ✓ Calcium and/or vitamin D
  - ✓ Bisphosphonates, e.g., alendronate (Fosamax), risedronate (Actonel)
  - ✓ Other hormones, e.g., Teriparatide (Forteo)

- **Cost: ranges from $4 to $130 per month**

Where is the evidence?

Existing Evidence on the Treatment of Osteoporosis

14 Cochrane systematic reviews

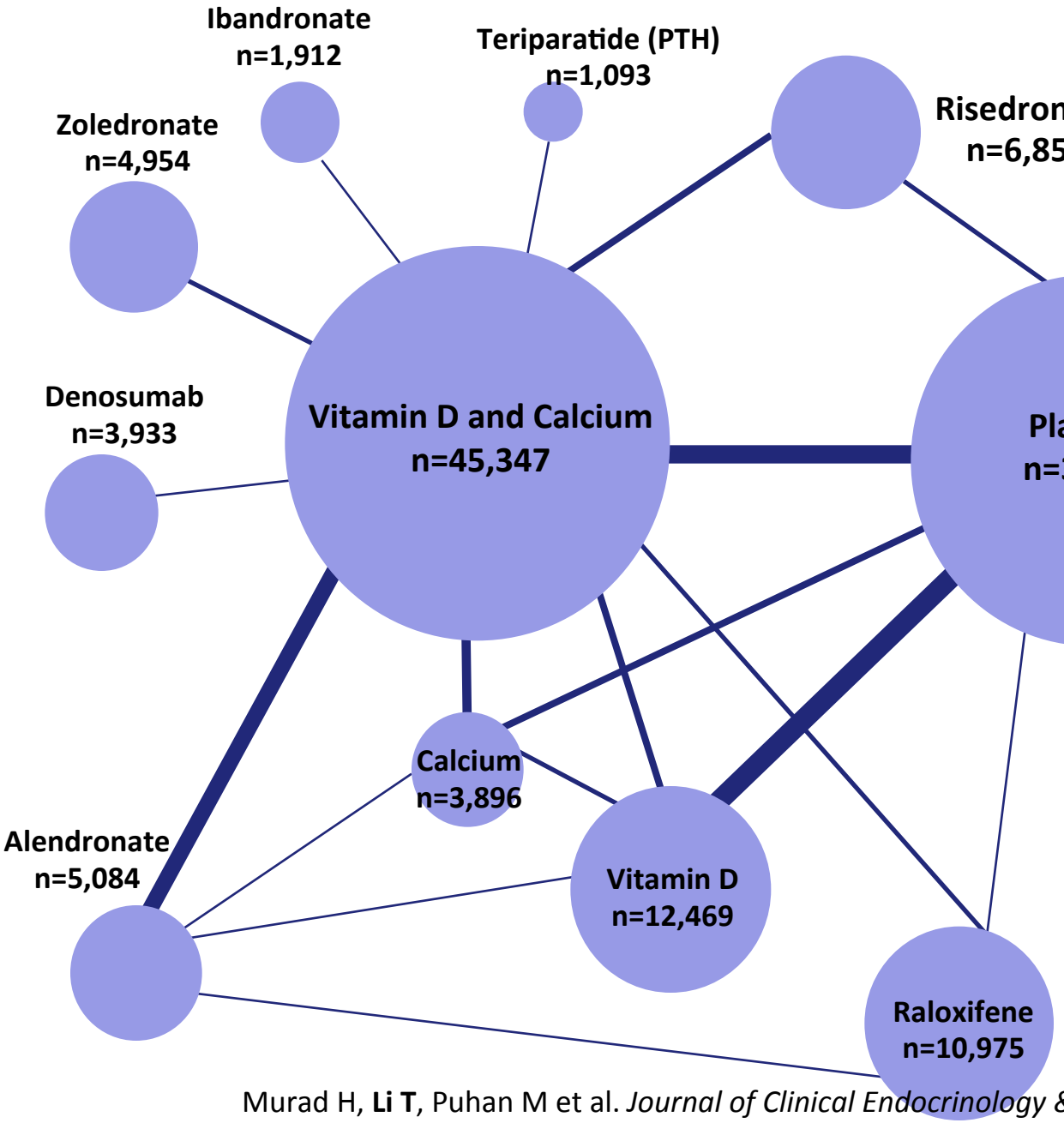
Which interventions work? In Whom?

*"At a dose of 10 mg per day, **alendronate** results in **a statistically significant and clinically important reduction** in vertebral, non-vertebral, hip and wrist fractures (Wells 2010)."*

*"**No statistically significant reductions** in non-vertebral, hip, or wrist fractures were found, regardless of whether **etidronate** was used for primary or secondary prevention (Wells 2010)."*

*"**Vitamin D alone** appears **unlikely to be effective in preventing hip fracture**…**Vitamin D with calcium** **reduces hip fractures** (Avenell 2009)."*

Treatment of Osteoporosis and Risk of Hip Fracture – Evidence Network



# of trials =39
# of participants =136,452
# of hip fracture =2561

KER UNIT
Mayo Knowledge Encounter Research Unit ◆ Mayo Clinic College of Medicine

Murad H, **Li T**, Puhan M et al. *Journal of Clinical Endocrinology & Metabolism.* 2012

What is a Network Meta-analysis?

Network (multiple treatments comparison) meta-analysis:

Meta-analysis, in the context of a systematic review, in which three or more treatments have been compared using both direct and indirect evidence from several studies.

Bucher 1997; Caldwell 2005; Glenny 2005; Song 2003

Why Use a Network Meta-analysis?

- **For many clinical conditions several active interventions are available**
  - e.g., Over 17 drugs (5 classes) for lowering intraocular pressure in patients with primary open angle glaucoma
- **Head-to-head randomized controlled trials (comparison of two active interventions) may be**
  - Unavailable or insufficient
  - Inconclusive or unreliable
- **Conventional systematic reviews focus on pair-wise, direct comparisons of interventions**
  - Indirect evidence has not been routinely synthesized
  - Challenge to rank multiple interventions

Indirect Comparisons of Multiple Treatments

Trial			
1	A	B	
2	A	B	
3		B	C
4		B	C
5	A		C
6	A		C
7	A	B	C

- Want to compare A vs. B
  Direct evidence from trials 1, 2 and 7
  Indirect evidence from trials 3, 4, 5, 6 and 7
- Combining all "A" arms and comparing with all "B" arms destroys randomization
- Use indirect evidence of A vs. C and B vs. C comparisons as additional evidence to preserve randomization and within-study comparison

Indirect Comparisons of Multiple Treatments

		Treatment		
Study	Sample Size per arm	A	B	C
1	100	20	15	
2	100	17	17	
3	50		9	6
4	50		7	4
5	200	18		16
6	200	19		13
7	500	8	7	5

Direct Comparison

- Use direct comparisons of A vs. B from trials 1, 2 and 7
- Uses only part of available information
- Gives valid, if less precise, estimate of treatment effect.
- Weight each study by its sample size

Study	Sample Size per arm	A	B
1	100	20	15
2	100	17	17
7	500	8	7

Direct estimate T_{AB}= (100*5 + 100*0 + 500*1) /700 = 1.43.

Indirect Comparison

- Use trials 3-6, contrasting the BC effect from trials 3 and 4 with the AC effect from trials 5 and 6

- Trial 7 also contributes to both parts of indirect comparison because 3-arm study estimates both AC and BC.

- But these two treatment comparisons include a common control treatment and so are correlated

Study	Sample Size per arm	A	B	C
3	50		9	6
4	50		7	4
5	200	18		16
6	200	19		13
7	500	8	7	5

Indirect Comparison

- For simplicity, ignore trial 7 in computing indirect comparison

Study	Sample Size per arm	A	B	C
3	50		9	6
4	50		7	4
5	200	18		16
6	200	19		13

Indirect estimate $T_{AC} - T_{BC}$

$= \{(200*2+200*6)/400 - (50*3+50*3)/100 = 1.0$

- Uses data from more studies but less efficiently
- Assumes similarity of comparisons

Naive Comparison

- Combine results from A arms in trials 1, 2, 5, 6 and 7 and compare their average to those from combining B arms in trials 1, 2, 3, 4 and 7

Study	Sample Size per arm	A	B
1	100	20	15
2	100	17	17
3	50		9
4	50		7
5	200	18	
6	200	19	
7	500	8	7

Naive Comparison

- Does not respect randomization within each study
- Studies with treatment A generally larger in size and outcome than those with B

Study	Sample Size per arm	A	B
1	100	20	15
2	100	17	17
3	50		9
4	50		7
5	200	18	
6	200	19	
7	500	8	7

Naïve estimate

(100*20+100*17+200*18+200*19*+500*8)/1100 -

(100*15+100*17+50*9+50*7+500*7)/800

= 4.34

Naive Methods – Pooling Study Arms across Trials

"When looking at all the study arms of either timolol or the lipid class drugs…"

Table 3. *Efficacy of IOP-lowering drugs (all studies)*

	Timolol	Latanoprost	Latanoprost + timolol	Bimatoprost	Bimatoprost + brimonidine
0–6 months data					
No. of study arms	21*	33†	11‡	18§	1**
No. of completed patients after 6 months	1946	2135	746	2326	13
Baseline IOP (mmHg), weighted mean	25.62	24.84	24.72	25.74	24.80
IOP reduction (mmHg), mean	5.19	6.44	5.85	7.13	8.50
IOP reduction (mmHg), weighted mean	5.78	6.69	6.18	7.81	8.50
IOP%–reduction, weighted mean	22.2%	26.7%	24.1%	30.3%	34.3%

Excerpted from:
Holmstrom et al. Curr Med Res Opin 2005: 21(11) 1875-1833

Mixed Comparison

- Use data from all 7 trials
- Combine direct and indirect estimates in a mixed treatment comparison
- 3-arm trial here introduces correlation between AC and BC and provides direct AB comparison, which is also correlated with the AC and BC comparisons
- Weighting by inverse variance of each estimate gives more weight to direct estimate because its variance is ¼ of variance of indirect estimate based on same amount of data

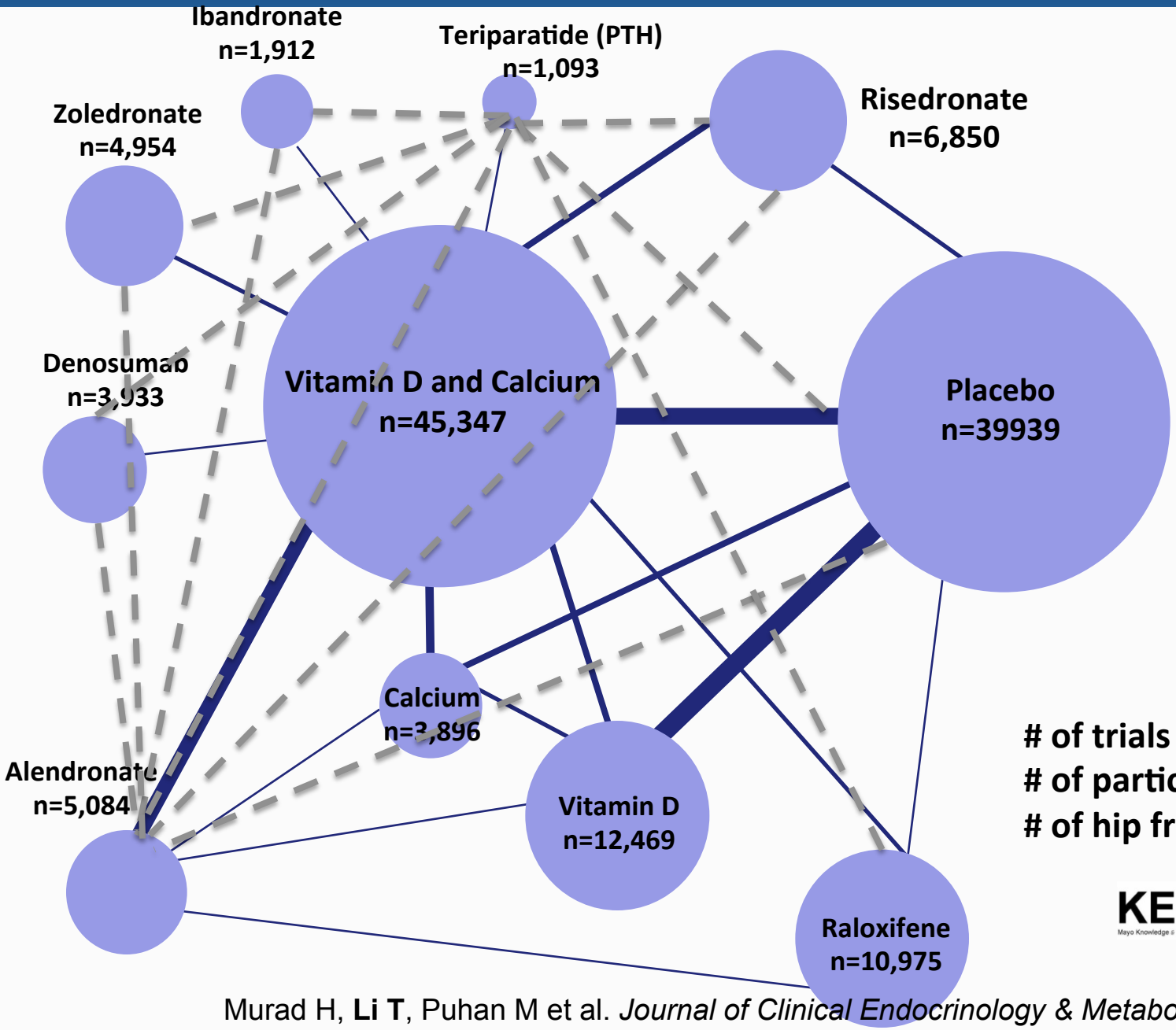
Mixed Comparison

		Treatment		
Study	Sample Size per arm	A	B	C
1	100	20	15	
2	100	17	17	
3	50		9	6
4	50		7	4
5	200	18		16
6	200	19		13
7	500	8	7	5

- Assume trial 7 only provides direct estimate
- Weight by total sample sizes

Mixed estimate = {(700*1.43 + 500*1.0)/1200 = 1.25

Treatment of Osteoporosis and Risk of Hip Fracture – Evidence Network



# of trials =39
# of participants =136,452
# of hip fracture =2561

KER UNIT
Mayo Knowledge Encounter Research Unit ◆ Mayo Clinic College of Medicine

Murad H, **Li T**, Puhan M et al. *Journal of Clinical Endocrinology & Metabolism (in press)*

Treatment of Osteoporosis and the Risk of Hip Fracture

Table. Pair-wise odds ratio and 95% credible interval

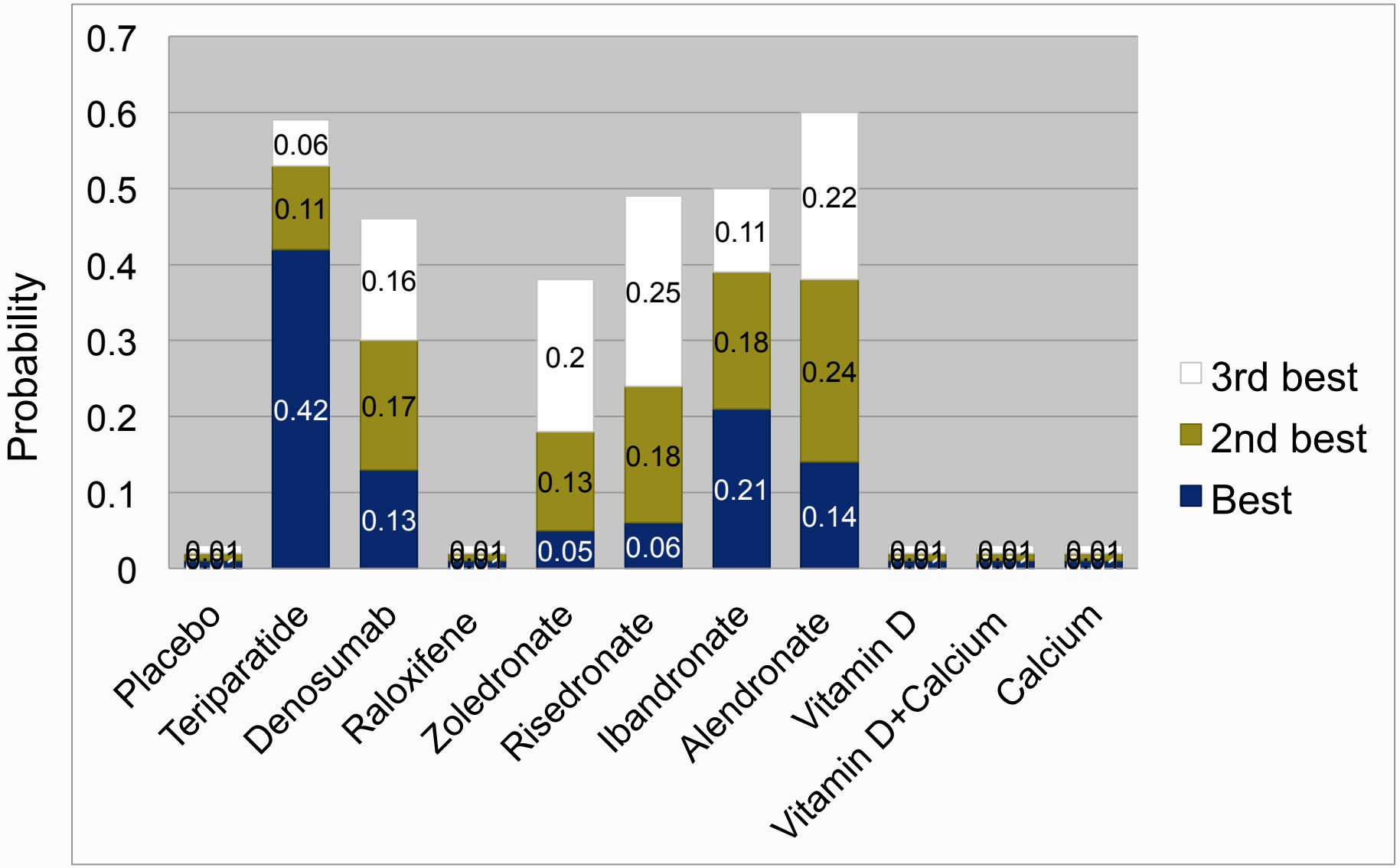
- Odds ratio <1 favors the treatment in the row
- Odds ratio >1 favors the treatment in the column
- # of trials =39
- # of participants =136,452
- # of hip fracture =3,850

Zoledronate		
0.97 (0.55; 1.51)	**Risedronate**	
0.94 (0.38; 2.44)	0.97 (0.41; 2.55)	**Ibandronate**
0.90 (0.52; 1.52)	0.93 (0.54; 1.60)	0.94 (0.36; 2.41)

Murad H, **Li T**, Puhan M et al. *Journal of Clinical Endocrinology & Metabolism (in press)*

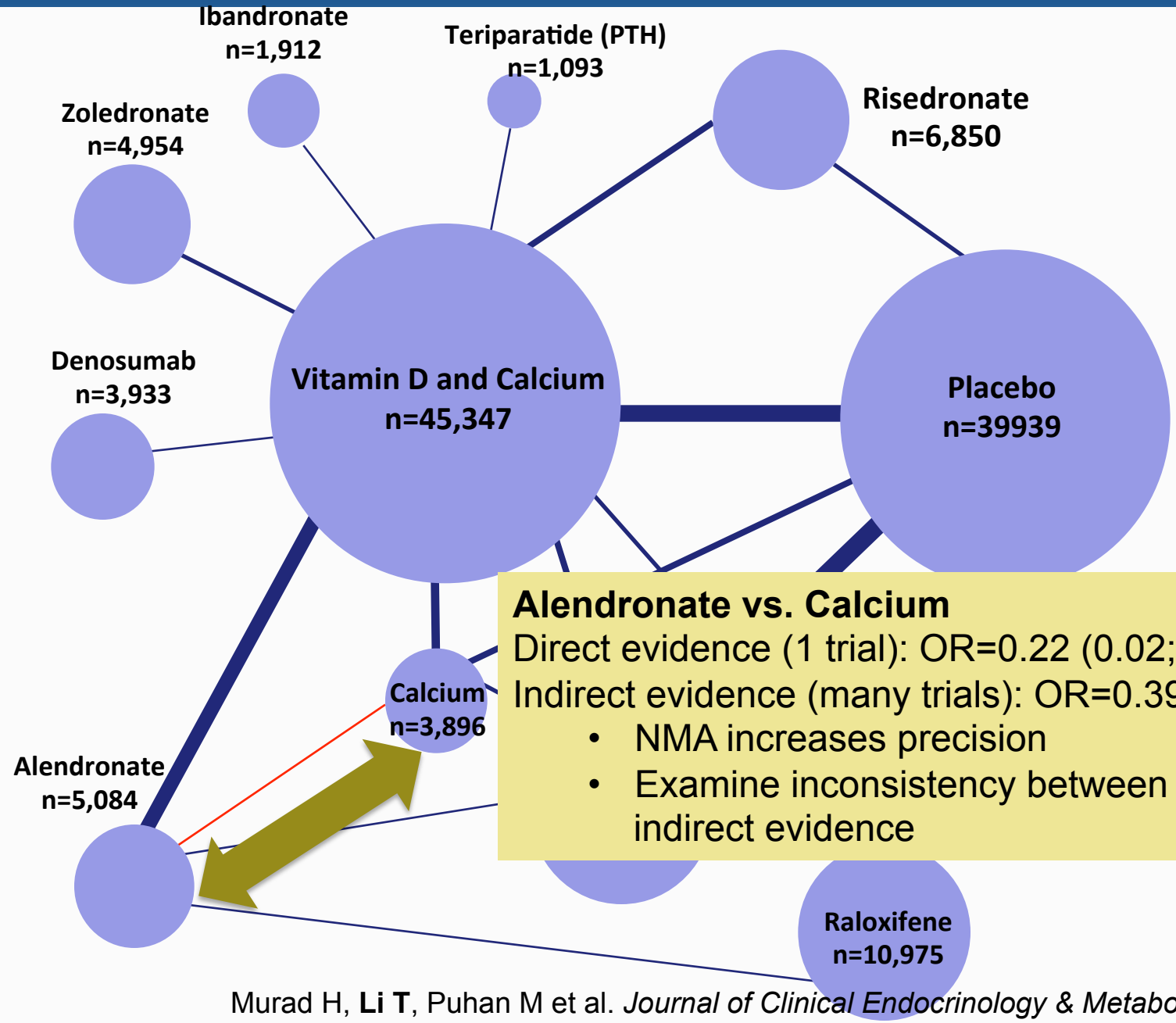
KER UNIT
Mayo Knowledge & Encounter Research Unit ◆ Mayo Clinic College of Medicine

Probability Ranking of Drugs in Reducing the Risk of Hip Fracture



Murad H, Li T, Puhan M et al. J Clin Endocrinol Metab. 2012;97(6):1871-80.

Treatment of Osteoporosis and Risk of Hip Fracture – Evidence Network



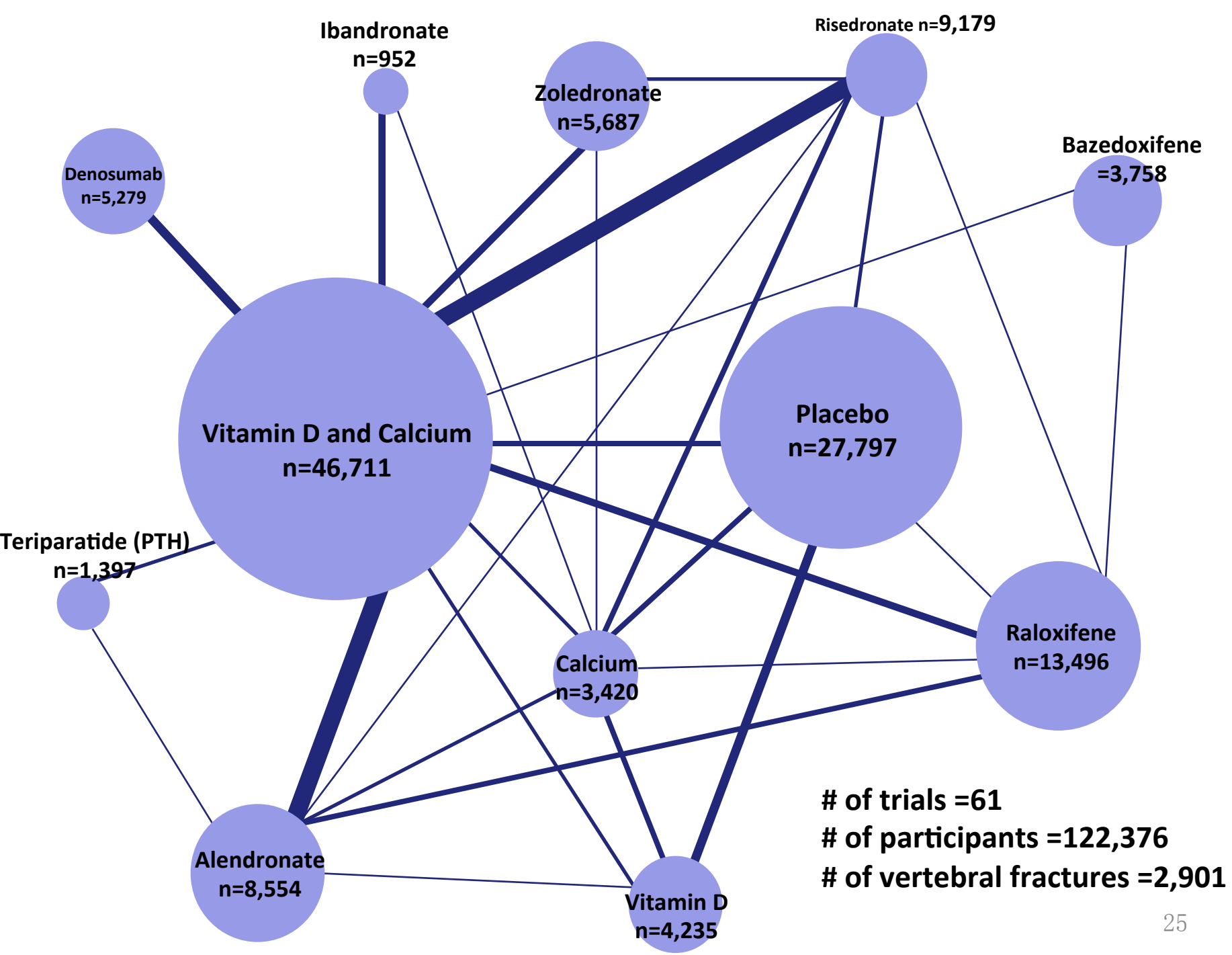
Alendronate vs. Calcium

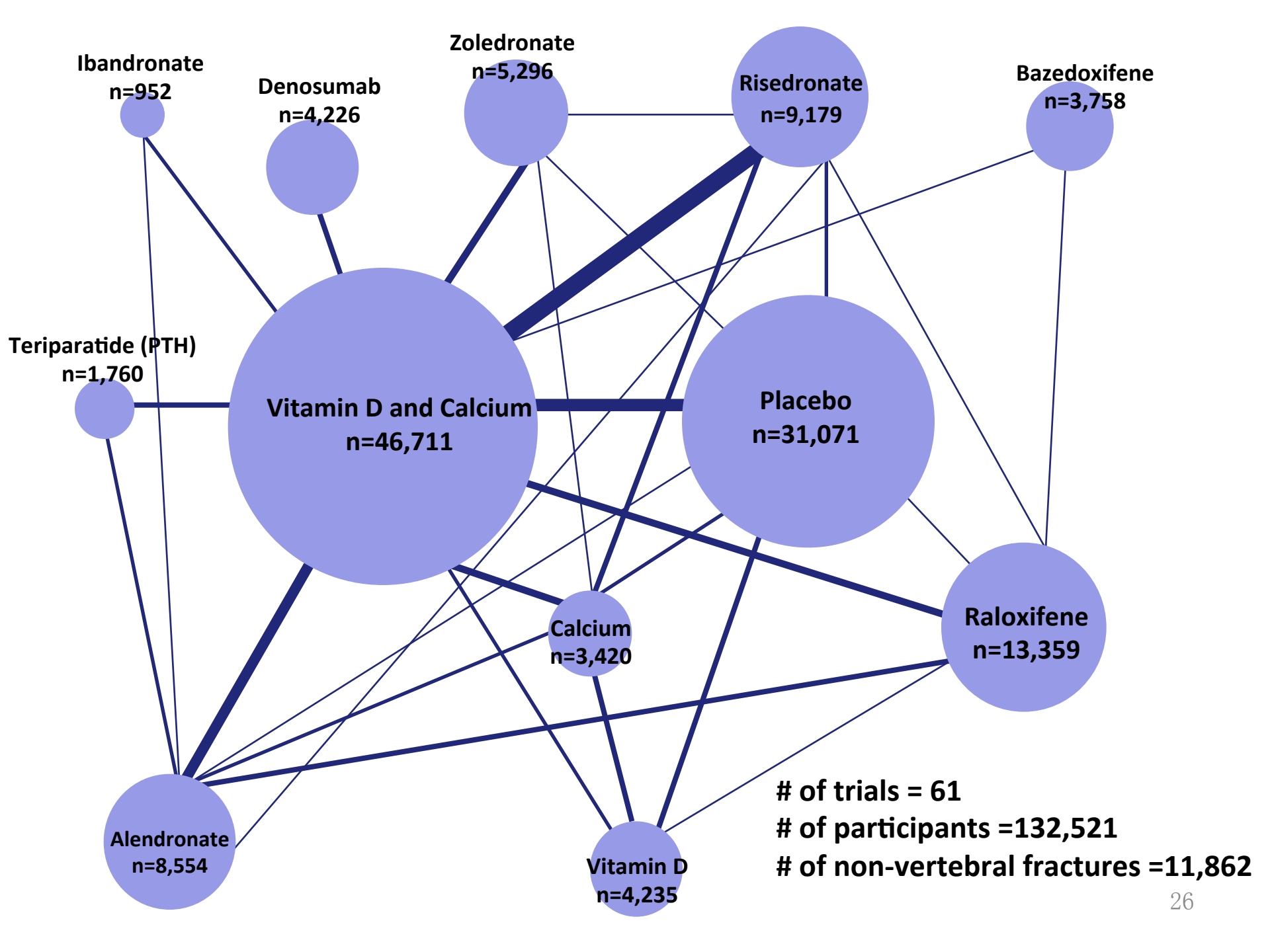
Direct evidence (1 trial): OR=0.22 (0.02; 2.13)

Indirect evidence (many trials): OR=0.39 (0.23; 0.65)

- NMA increases precision
- Examine inconsistency between direct and indirect evidence

Murad H, **Li T**, Puhan M et al. *Journal of Clinical Endocrinology & Metabolism (in press)*


Ibandronate
n=952
Risedronate n=9,179
Zoledronate
n=5,687
Denosumab
n=5,279
Bazedoxifene
=3,758
Vitamin D and Calcium
n=46,711
Placebo
n=27,797
Teriparatide (PTH)
n=1,397
Raloxifene
n=13,496
Calcium
n=3,420
Alendronate
n=8,554
Vitamin D
n=4,235
# of trials =61
# of participants =122,376
# of vertebral fractures =2,901


Ibandronate
n=952
Denosumab
n=4,226
Zoledronate
n=5,296
Risedronate
n=9,179
Bazedoxifene
n=3,758
Teriparatide (PTH)
n=1,760
Vitamin D and Calcium
n=46,711
Placebo
n=31,071
Calcium
n=3,420
Raloxifene
n=13,359
Alendronate
n=8,554
Vitamin D
n=4,235
# of trials = 61
# of participants =132,521
# of non-vertebral fractures =11,862

Basic Assumptions

- ✓ **Transitivity**

  Trials involving treatments needed to make indirect comparisons are comparable so that it makes sense to combine them

  Needed for valid indirect comparison estimates

- ✓ **Consistency**

  Direct and indirect estimates give same answer

  Needed for valid network meta-analysis estimates

Check the Assumptions for Analysis

Example 1: Consider a placebo that may be given in an oral or an intravenous form.

If treatment A is an oral treatment and treatment B is an intravenous one, then it may not be valid to compare A and B indirectly through the placebo C if the different routes of administration produce different effects.

This may violate the transitivity assumption because…

Five Interpretations of Transitivity

1. Participants included in the network could in principle be randomized to any of the three treatments A, B, C.
2. Treatment C is similar when it appears in AC and BC trials
3. 'Missing' treatment in each trial is missing at random
4. There are no differences between observed and unobserved relative effects of AC and BC beyond what can be explained by heterogeneity
5. The two sets of trials AC and BC do not differ with respect to the distribution of effect modifiers

Salanti (2012)

Violate the Transitivity Assumption

Example 1: Consider a placebo that may be given in an oral or an intravenous form.

1. The different protocols would preclude examining all treatments together in the same study;
2. The placebo has a different route of administration in the two types of trials;
3. The treatment omitted is not given because it requires a different protocol;
4. The unobserved treatment effect might come from a different distribution than the one observed because it would have a different mode of administration;
5. The route of administration is a potential effect modifier of the treatment effect.

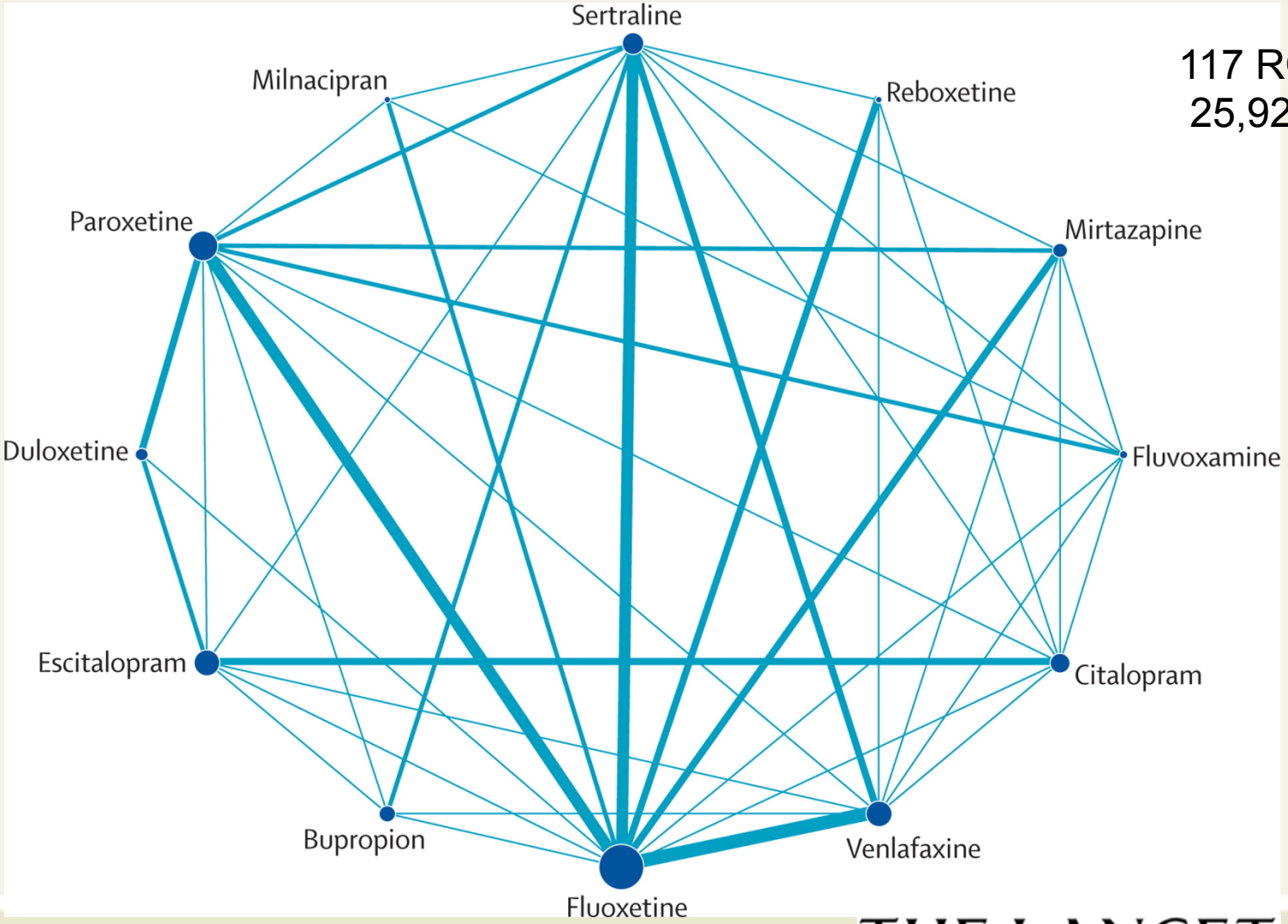
Check the Assumptions for Analysis

Example 2: Intervention A is clinically indicated only for previously untreated patients and intervention B is clinically indicated only when all other treatments have failed.

- ✓Initial interventions (for treatment naïve patients) and add-on interventions could be studied in the same review.
- ✓The key is to analyze incomparable interventions and distinct populations in separate network meta-analyses.

Methodological Challenges and Research Opportunities for Network Meta-analysis

Evidence Network of Comparative Efficacy and Acceptability of 12 New Generation Antidepressants



117 RCTs
25,928 participants

Cipriani et al. Lancet 2009; 373:746-58

THE LANCET

Potential Bias in Study and Data Selection - Publication Bias

- "Among placebo-controlled antidepressant trials registered with the FDA, ***most negative results are unpublished or published as positive.***"
  - 5 sertraline trials registered with FDA
    - 1/5 positive trial was published
    - 1/5 negative trial was published as positive
    - 3/5 were never published

Correspondence: Ioannidis JP. Lancet 2009; 373:1759-1760

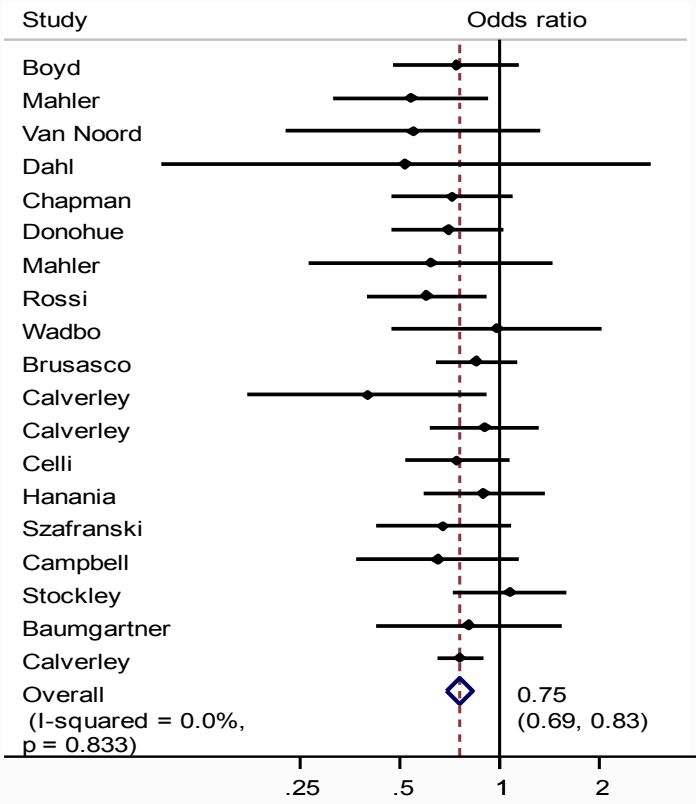
Potential Bias in Study and Data Selection - Publication Bias (cont'd)

Discrepant Rankings of Effect Sizes for Effectiveness of Antidepressants

	Published placebo-controlled trials[2]	Registered placebo-controlled trials[2*]	Head-to-head comparisons[1]	Both placebo-controlled and head-to-head[3]
Bupropion	12	12	6	3
Citalopram	10	11	5	11
Duloxetine	7	7	10	7–8
Escitalopram	8–9	6	2	9
Fluoxetine	11	8–10	9	1
Mirtazapine	2	4	1	7–8
Nefazodone	5	8–10	..	4
Paroxetine	1	1	7	10
Paroxetine CR	8–9	5	†	†
Sertraline	6	8–10	4	2
Venlafaxine	3–4	2–3	3	6
Venlafaxine XR	3–4	2–3	‡	‡
Milnacipran	..	..	8	..
Fluvoxamine	..	..	11	12
Reboxetine	..	..	12	..
Trazodone	..	..	..	5

Correspondence: Ioannidis JP. Lancet 2009; 373:1759-1760

Conventional Meta-analysis: Entire Evidence for 1 Estimate



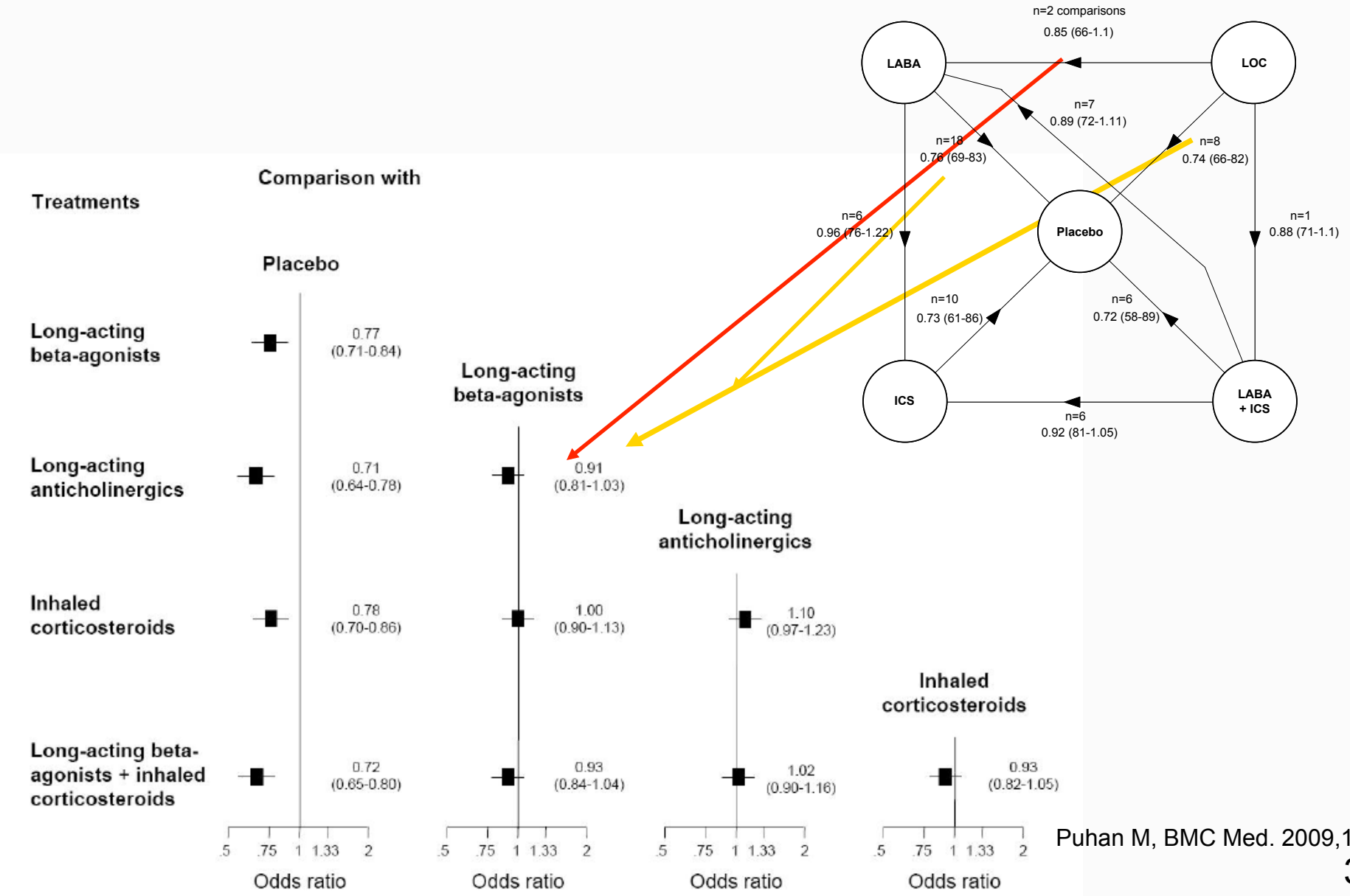
X trials inform

1 point estimate

Quality of evidence

- Risk of bias (Cochrane)
- Summary of quality items
- ●●●○ (GRADE)

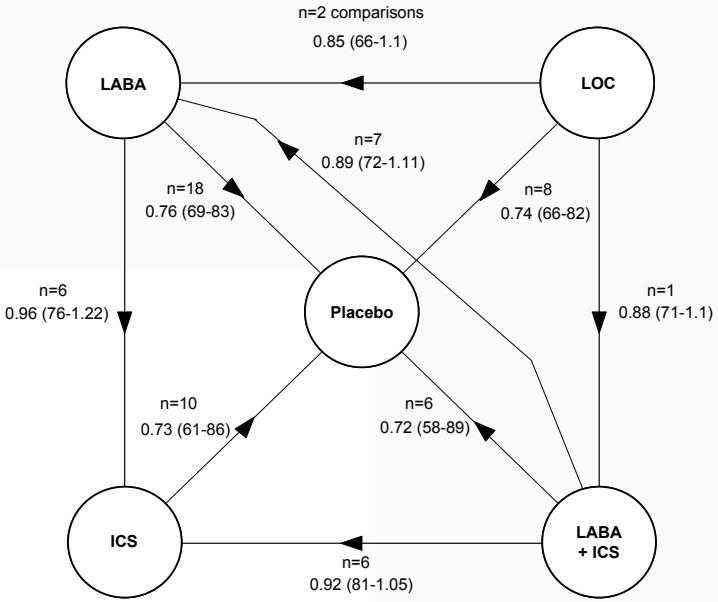
Network meta-analysis: Trials contribute to different estimates

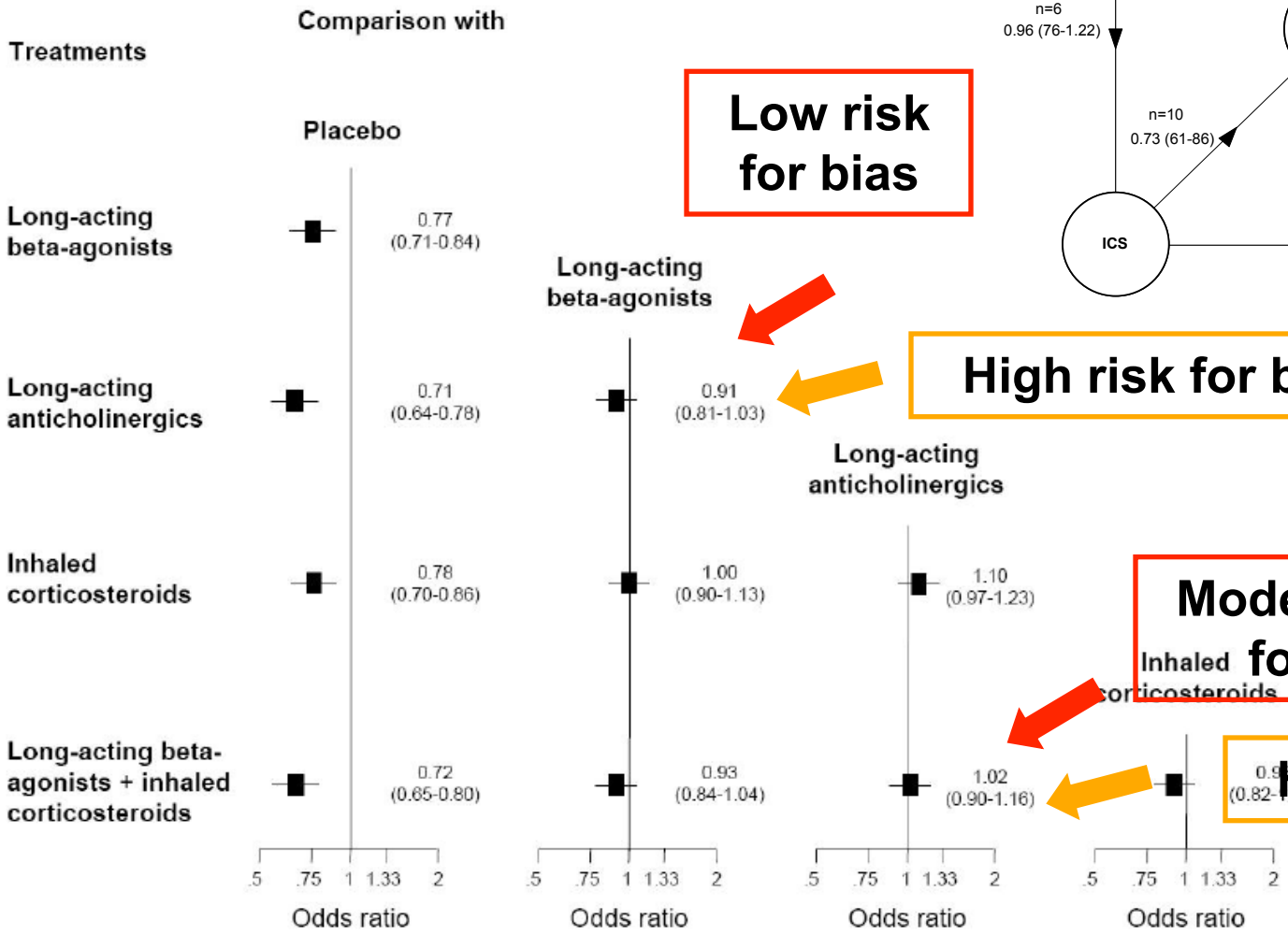


Puhan M, BMC Med. 2009,14;7:2.

Quality of evidence likely to be heterogeneous across network

Within and across comparisons





Low risk for bias

High risk for bias

Moderate risk for bias

High risk for bias

Should This Review Be an Overview or an Intervention Review?

- Misconceptions
  - Any review that compares interventions must use the Overview format
  - Any review that compares multiple interventions must include indirect comparisons

Differences Between Overviews and Intervention Reviews

- **Search Strategy**
  - Intervention reviews search for trials
  - Overviews search for reviews
- **Approach to Analysis**
  - Intervention reviews use a trial level analysis
  - Overviews may be able to use a review level analysis

- **Search Strategy** – Always differs
  - Intervention reviews search for trials
  - Overviews search for reviews
- **Approach to Analysis** – ~~Sometimes differs~~
  - Intervention reviews use a trial level analysis
  - Overviews may be able to use a review level analysis
    - ~~But will often use a trial level analysis instead~~

- The Intervention Review format is ***strongly recommended*** for reviews that include indirect comparisons.
- Because these comparisons require detailed knowledge of the trials.

Where does this leave Overviews?

Some Overviews Do Not Compare Interventions

- Different outcomes of a single intervention
  - e.g. Hormone Replacement Therapy
- Different conditions, problems, or populations
  - e.g. Aspirin to prevent stroke
- Related non-competing interventions

Related Interventions - Liver Transplantation

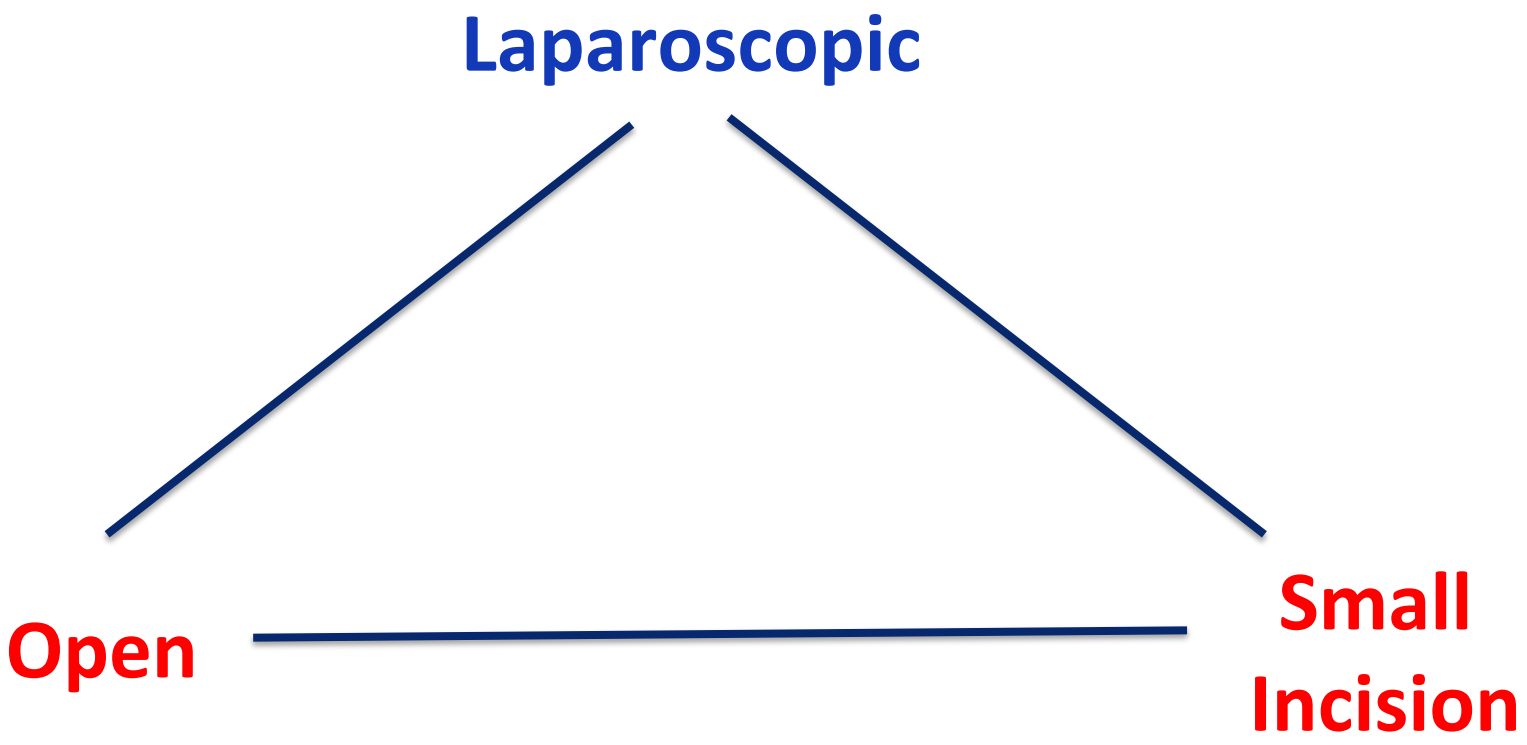
- Tasks of the transplant team
  - Remove the diseased liver
  - Remove the donor liver
  - Keep the donor liver viable
  - Transplant the donated liver
    - 5 different anastomoses to be made
- 9 relevant Cochrane reviews

Can Overviews Sometimes Be Used to Compare Interventions?

- Still under active discussion
- Your input needed

- Some examples from existing overviews
  - Indirect comparisons based on summary statistics from the review
  - Direct comparisons only
  - Analogous comparisons

Surgical Techniques for Cholecystectomy



Direct comparisons from 3 Cochrane Reviews

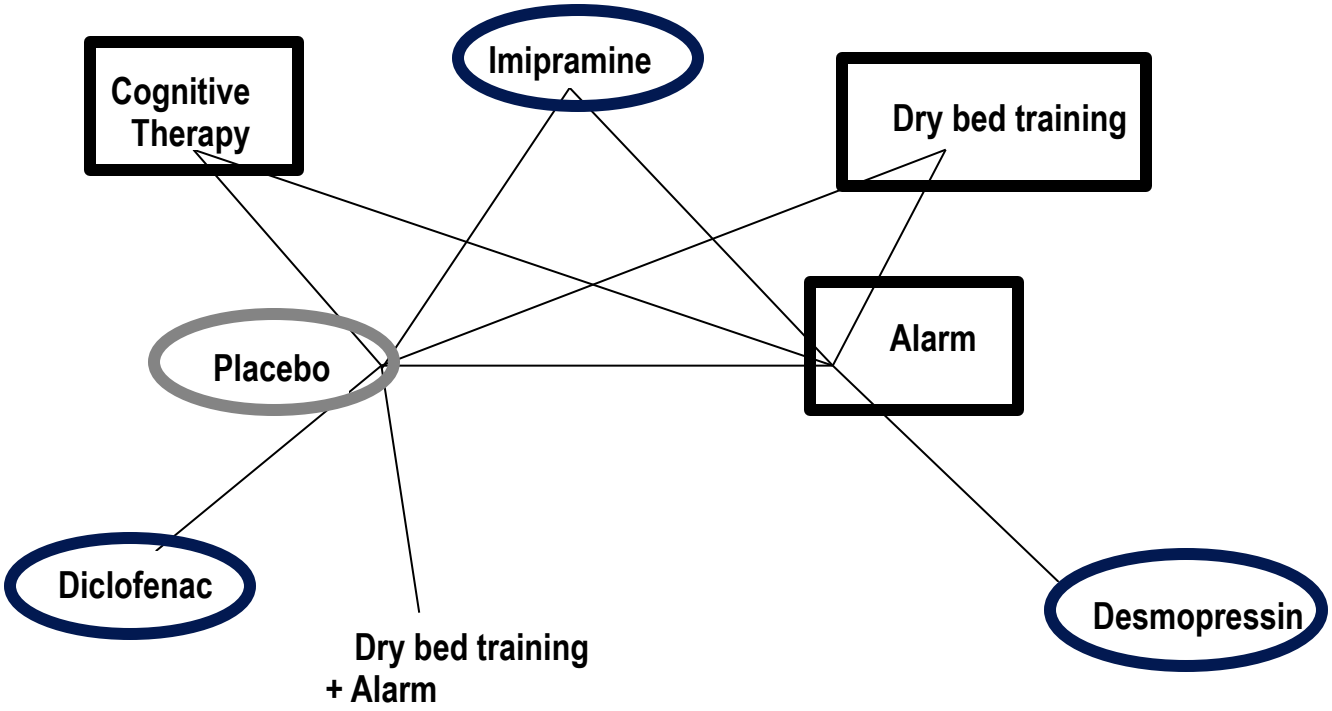
http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008318/abstract

Direct Only - Surgical Techniques for Laparoscopy

- Simple network
- All direct comparisons covered by existing Cochrane Intervention Reviews
- All reviews were up to date
- Overview authors were also the authors of all 3 Intervention Reviews

- How often are all of these conditions met?
- When is the direct evidence "good enough" on its own?

Interventions for Enuresis



J Clin Epidemiol. 63:875-82 PMID: 20080027

Trial or Review summaries for NMA?

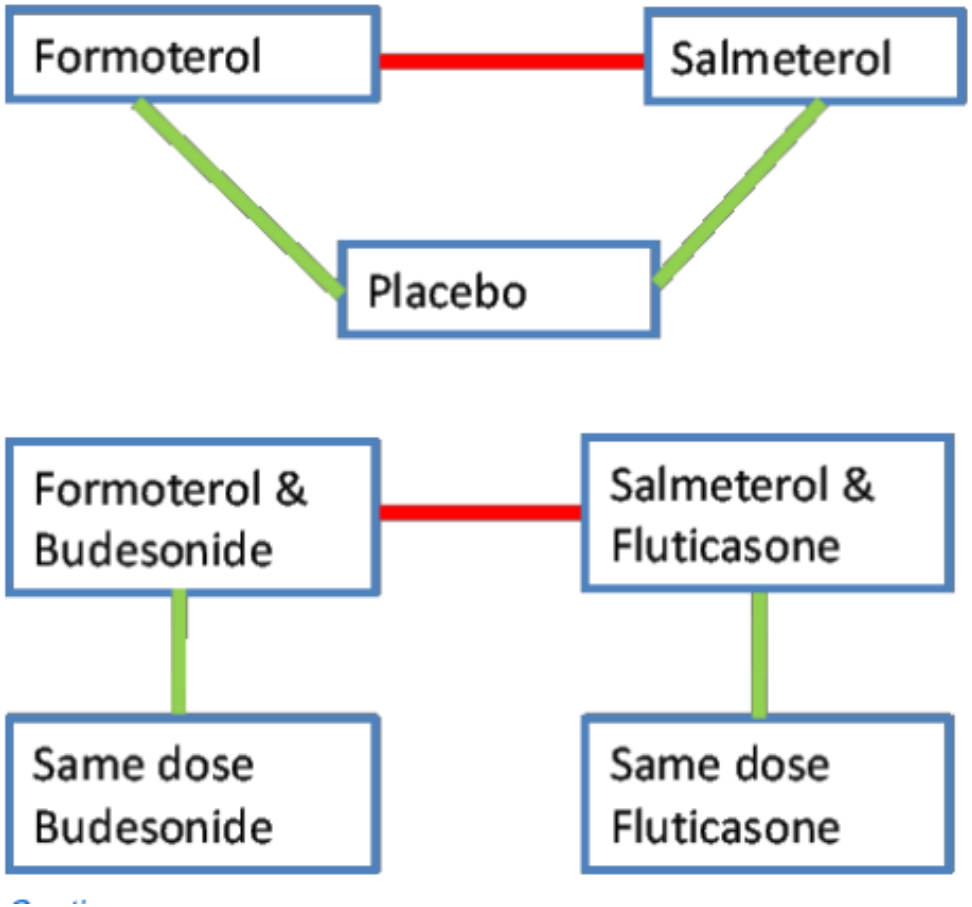
	Review level summaries		Trial level summaries	
Treatment	***Prob best***	***RR (no treatment)***	***Prob best***	***RR (no treatment)***
No treatment	0	1	0	1
Alarm	0.08	0.40 *(0.31, 0.53)*	0.03	0.41 *(0.30, 0.53)*
DBT	0	0.82 *(0.66, 1.03)*	0.01	0.82 *(0.66, 1.02)*
Desmopressin	0	0.54 *(0.35, 0.84)*	0.04	0.58 *(0.37, 0.88)*
Imipramine	0	0.68 *(0.53, 0.89)*	0	0.69 *(0.52, 0.89)*
Psych therapy	0.01	0.65 *(0.35, 1.22)*	0.02	0.69 *(0.35, 1.22)*
DBT + alarm	0.78	0.19 *(0.05, 0.76)*	0.78	0.24 *(0.05, 0.73)*
Diclofenac	0.13	0.46 *(0.16, 1.38)*	0.12	0.53 *(0.16, 1.35)*

Work in progress: do not cite

A Possible Exception?

- It may be possible to do indirect comparisons using summary statistics from review meta-analyses.
- This doesn't happen very often.
- You still need to know the trials in detail.
- But review authors often know the trials well.

Analogous - Adverse effects of LABAs for Asthma



http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD010005/abstract

Conclusions

- In many problems, investigators would like to synthesize evidence from multiple interventions tested in multiple trials.
- When good trial-level data that satisfy assumptions of network meta-analysis are available and goal is to rank interventions, network meta-analysis is preferred.
- When the objective is not to compare competing interventions, network meta-analysis is not useful.

Key Messages

- Network meta-analysis is an extension of standard, pair-wise meta-analysis.
- Use of network meta-analysis is often necessary for drawing inference about multiple competing interventions and a formal approach is preferable.
- Intervention reviews are encouraged if indirect comparisons are to be performed.
- The choice between the Intervention Review or Overview format is less clear for reviews where no indirect comparisons are planned.