UNIVERSITY OF EXETER | MEDICAL SCHOOL

Heterogeneity in meta-analysis: What it is and what to do with it

Jaime Peters

j.peters@exeter.ac.uk

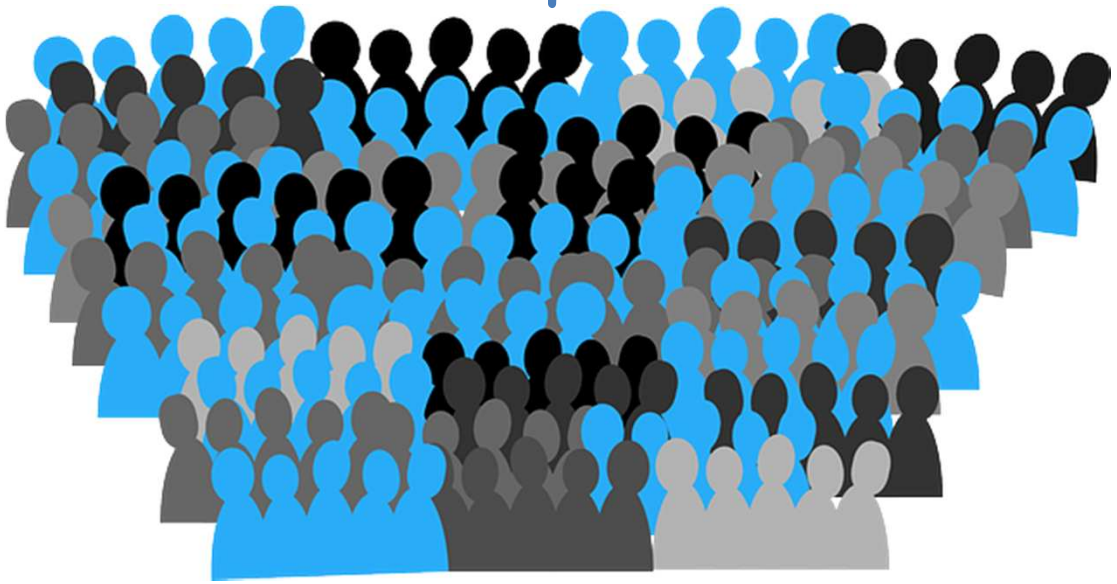
NIHR CLAHRC
South West Peninsula

NHS
National Institute for
Health Research





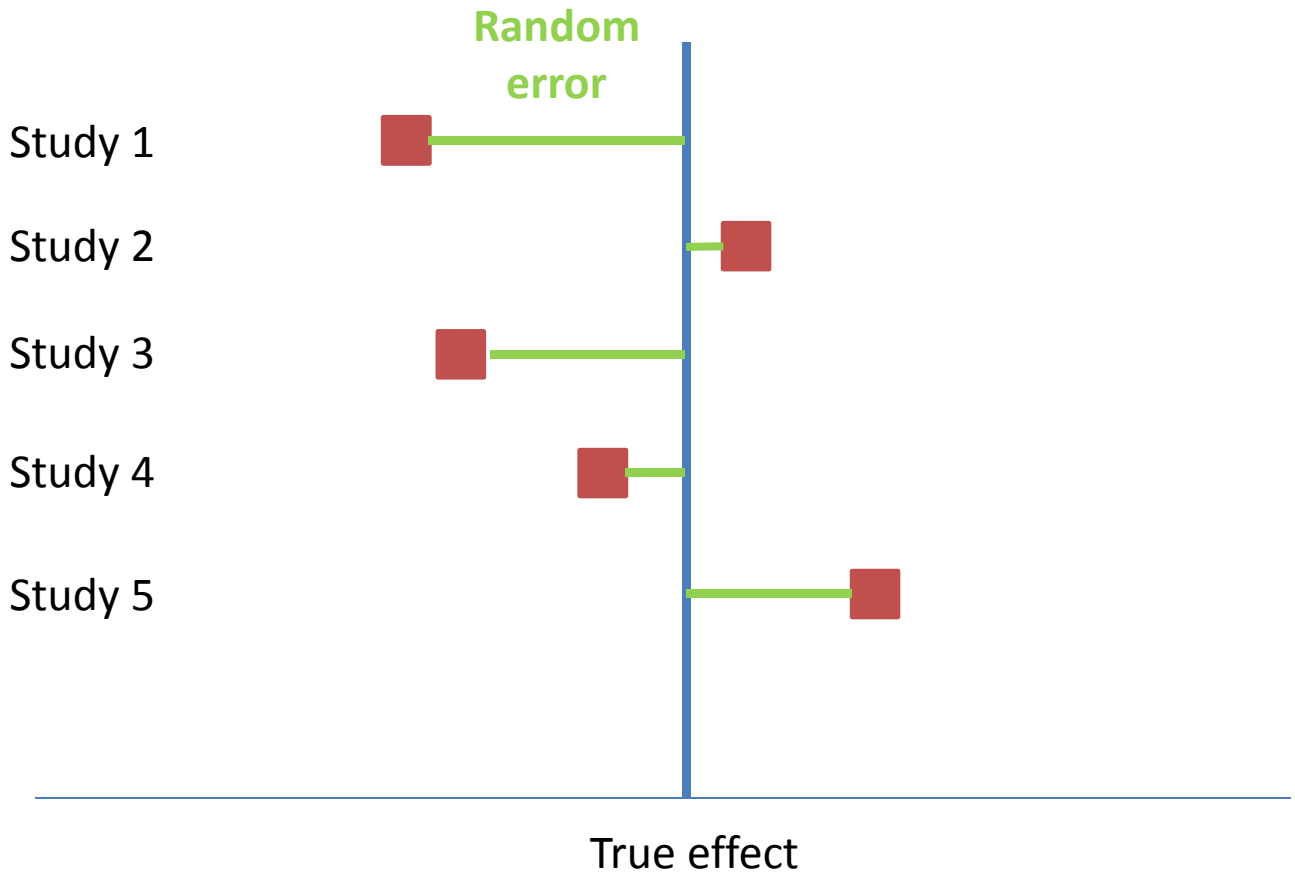
Weighted average of the results of each study



Fixed effects meta-analysis

- Assumes all studies are estimating the same underlying effect (OR, HR, RR, mean difference)
- The only variation between estimates from studies is due to random error

Fixed effects meta-analysis



Acknowledgement to Julian Higgins

Fixed effects inverse-variance weighted model

$$Y = \frac{(1/var_1 \times y_1) + (1/var_2 \times y_2) + \ \ldots \ + (1/var_{10} \times y_{10})}{1/var_1 + 1/var_2 + \cdots + 1/var_{10}}$$

This can be generalised as below where the total number of studies in a meta-analysis is k

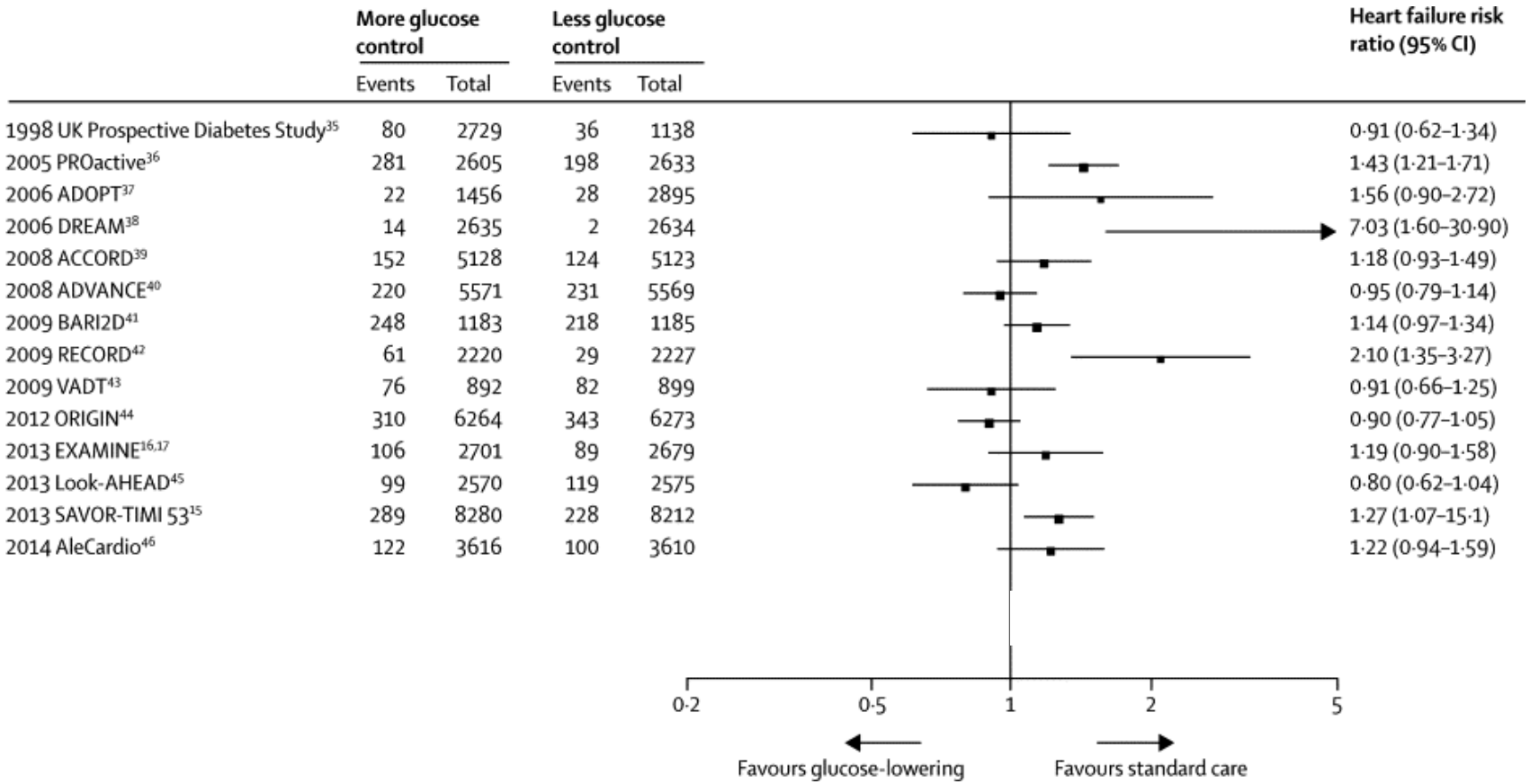
$$Y = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i} \qquad \text{where } w_i = 1/var_i$$

Fixed effects inverse-variance weighted model

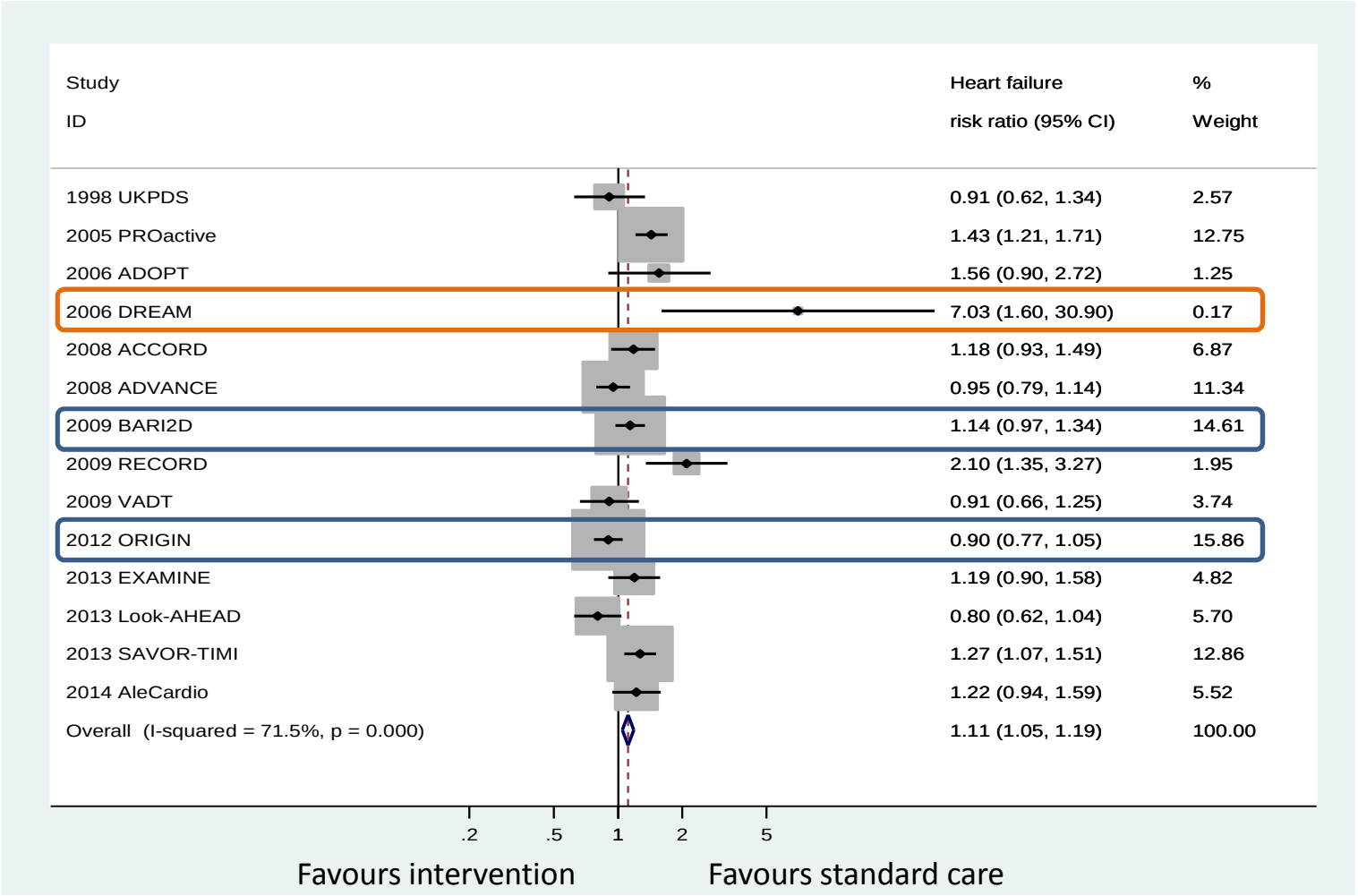
- Each study weighted by the inverse of their variance (precision)

- **Which of these studies would you expect to get the**
  - **least weight in the meta-analysis?**
  - **most weight in the meta-analysis?**

Forest plots



Fixed effects meta-analysis



Heterogeneity (between-study variation)

- Fixed effects model - only random error explains differences between studies
- However, rarely is it the case that the only differences between the effects across studies are due to random error
- Instead, there may be more heterogeneity (variability) between studies than is expected by chance
- **Why might study results differ?**

Heterogeneity
(between-study variation)

- Country
- Type of intervention
- How intervention is implementation
- Participant age
- Gender mix
- Improvements in technology, e.g. measuring outcomes
- Study design

- Clinical
- Methodological
- Statistical

Heterogeneity (between-study variation)

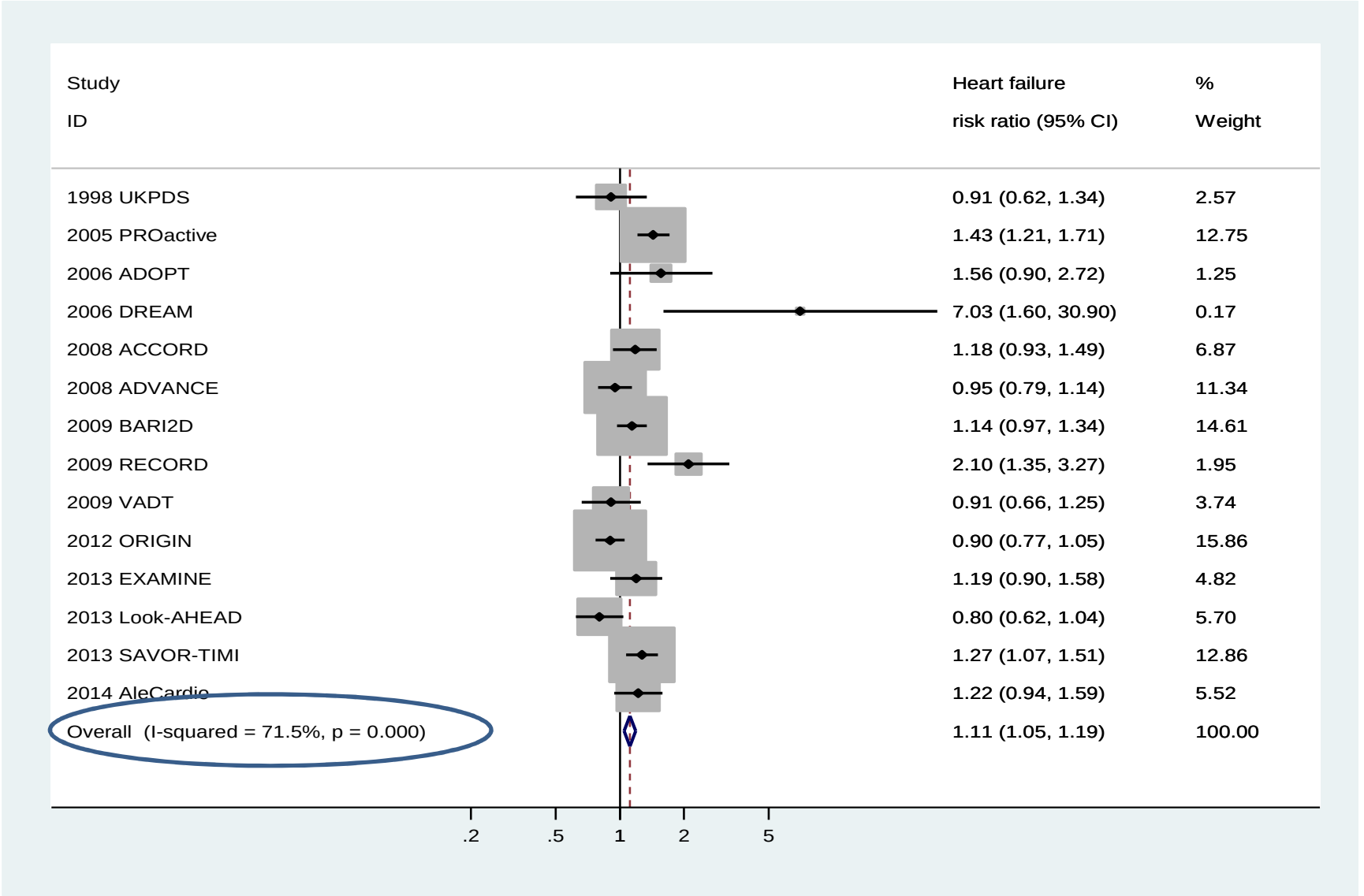
- Methods exist to
  - **Identify** whether heterogeneity exists
  - **Account** for the heterogeneity
  - **Explore** the sources of heterogeneity

- Understanding why there is heterogeneity between your study estimates may be more important than estimating some average effect

Identifying heterogeneity

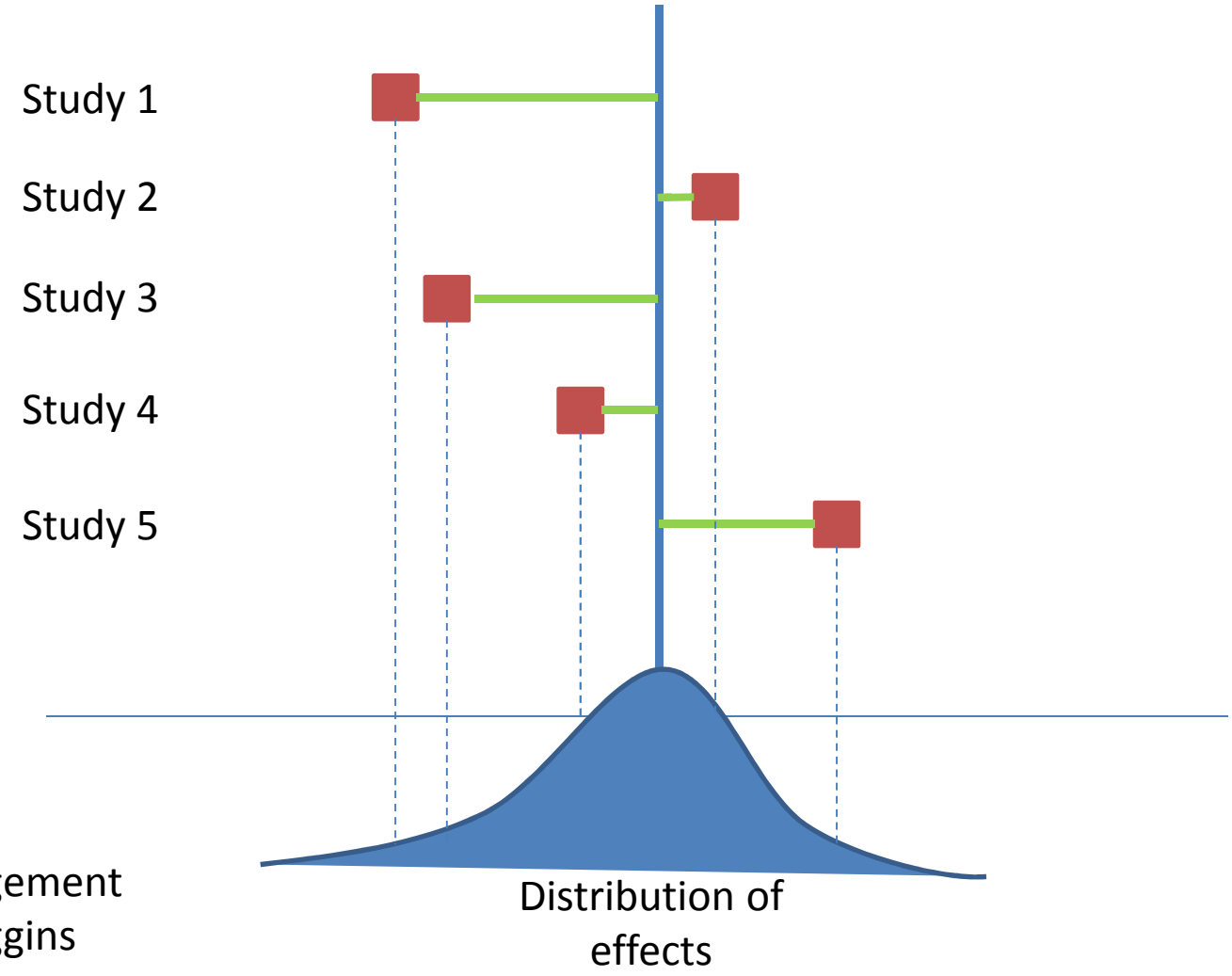
- χ^2 statistical test
  - tests that the true treatment effects are the same in all primary studies versus the alternative that at least one treatment effect differs from the others
  - Low statistical power to detect heterogeneity if it exists
  - Compare to significance level of 0.10 (rather than 0.5)
- I^2 (and 95% CI)
  - Measure of inconsistency across studies
  - Percentage of total variation across studies that is due to heterogeneity rather than chance
  - $0\% \leq I^2 \leq 100\%$; Low ≈ 25%, Moderate ≈ 50%, High ≈ 75%
  - Cochrane:
    - 0-40% might not be important
    - 30-60% may represent moderate heterogeneity
    - 50%-90% may represent substantial heterogeneity
    - 75%-100% considerable heterogeneity

Fixed effects meta-analysis



Accounting for heterogeneity

Random error PLUS estimate of between study variability, tau^2



Acknowledgement to Julian Higgins

Accounting for heterogeneity: Random effects inverse-variance weighted model

- In addition to weighting studies by the inverse of their variance, an estimate of between-study heterogeneity is also accounted for
- Using notation from above

$$Y = \frac{\sum_{i=1}^{k} w*_i y_i}{\sum_{i=1}^{k} w*_i} \qquad \text{where } w*_i = \frac{1}{var_i + tau^2}$$

and tau^2 = estimate of between-study heterogeneity

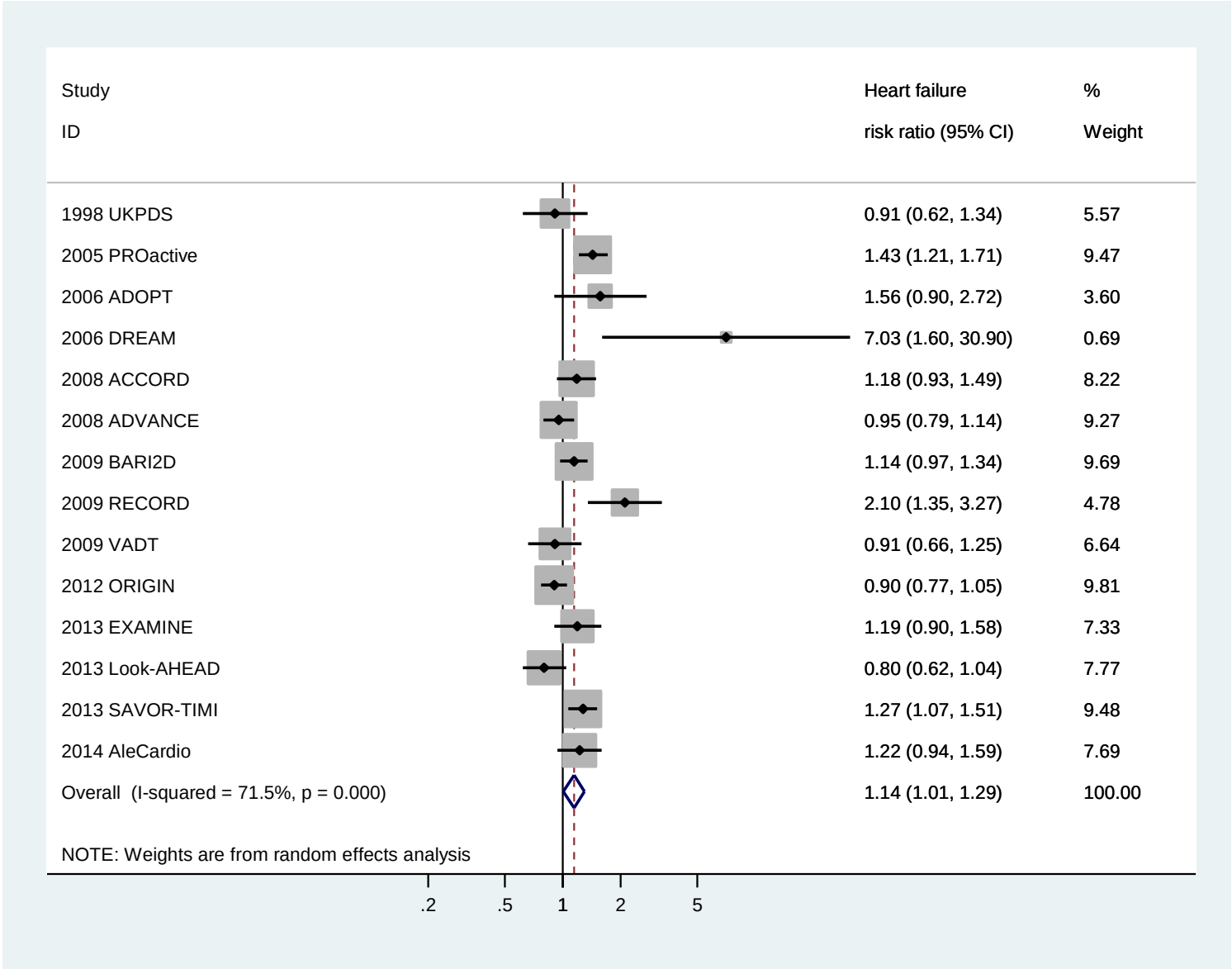
Random effects meta-analysis

$$\bar{Y} = \frac{(\frac{1}{var_1 + tau^2} \times y_1) + \ \ldots \ + (\frac{1}{var_{10} + tau^2} \times y_{10})}{\frac{1}{var_1 + tau^2} + \cdots + \frac{1}{var_{10} + tau^2}}$$

$$var(\bar{Y}) = \frac{1}{\sum_{i=1}^{k} w*_i} \qquad \text{where} \quad w*_i = \frac{1}{var_i + tau^2}$$

$$var(\bar{Y}) = \frac{1}{\frac{1}{var_1 + tau^2} + \cdots + \frac{1}{var_{10} + tau^2}}$$

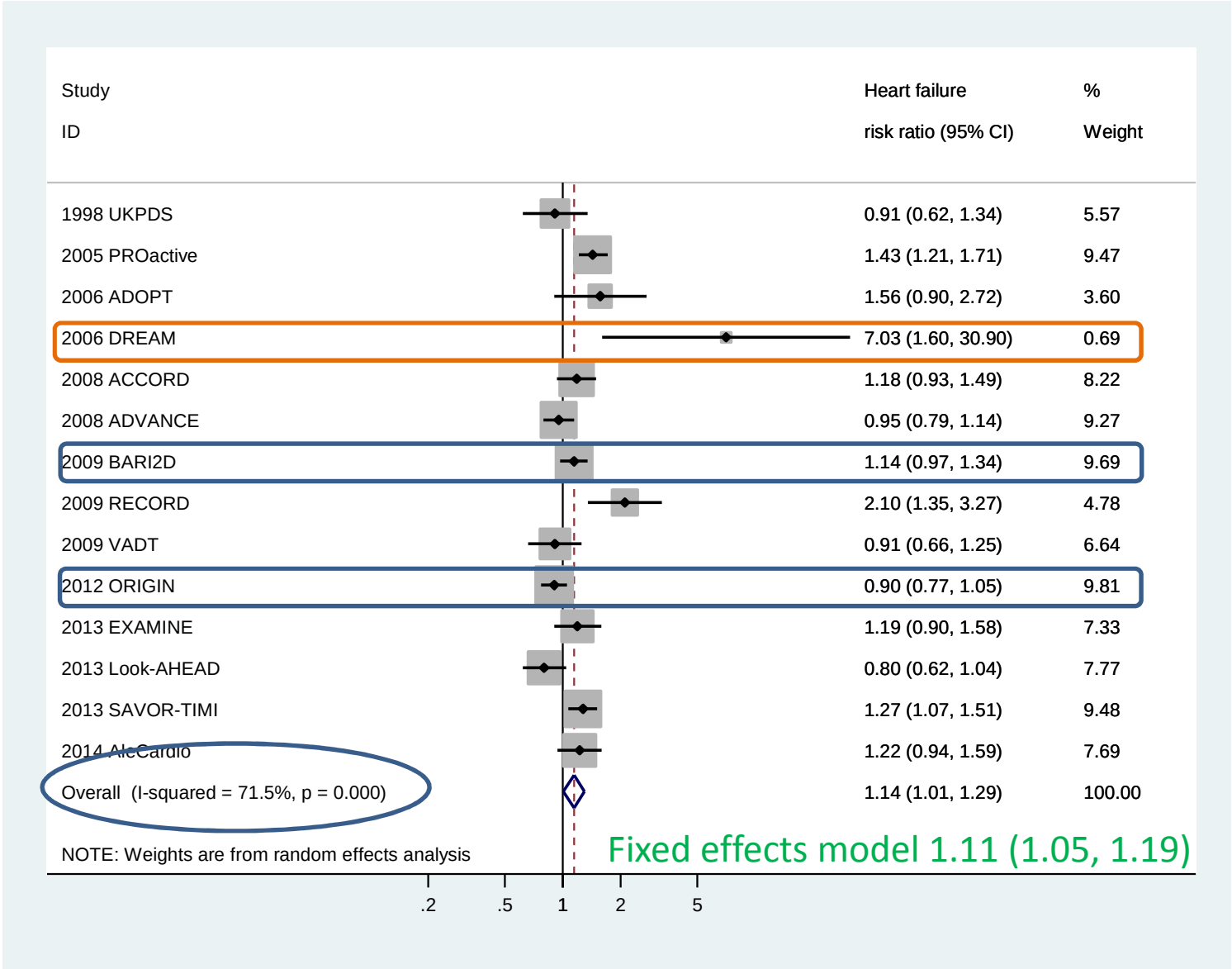
Random effects meta-analysis



Interpretation of random effects model

- **How does the random effects result differ to the fixed effects?**
- **What has happened to the study weights?**

Random effects meta-analysis



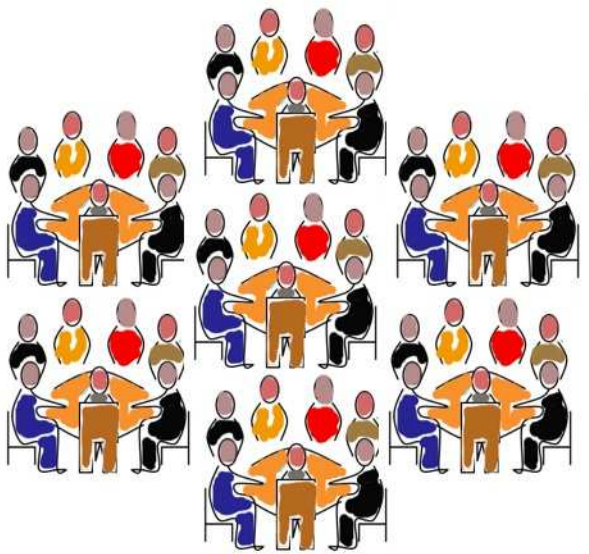
Fixed vs random effects?

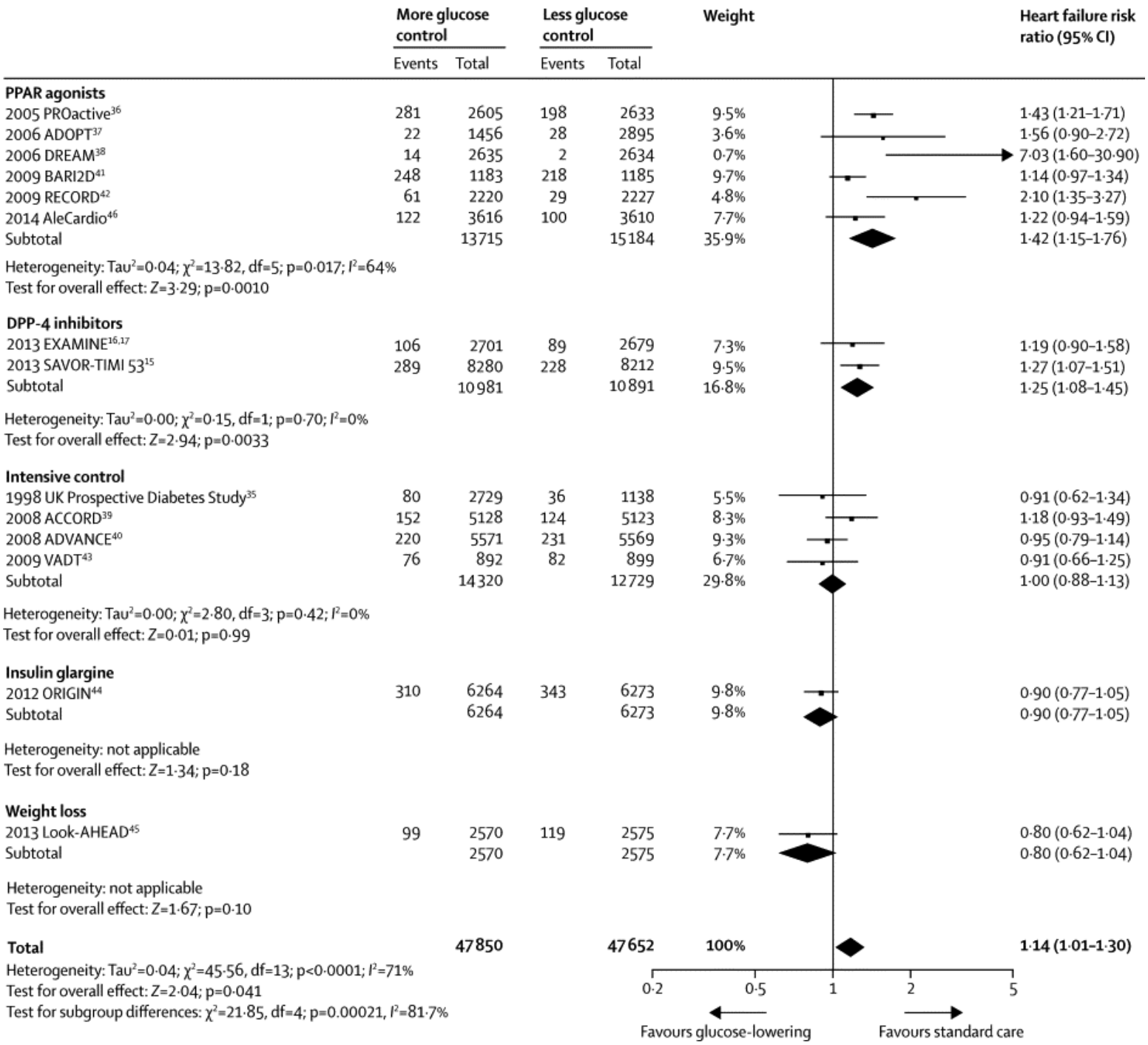
- No rules for when to use fixed vs random effects
- Random effects estimate incorporates the heterogeneity, but does not mean heterogeneity is no longer a problem

- Is the average effect useful?
- If too much heterogeneity, don’t do a meta-analysis of all studies
- **Understand why heterogeneity exists**

Exploring sources of heterogeneity (1)

- Subgroup analyses
  - By country
  - By type of intervention
  - By age group of participants
  - …



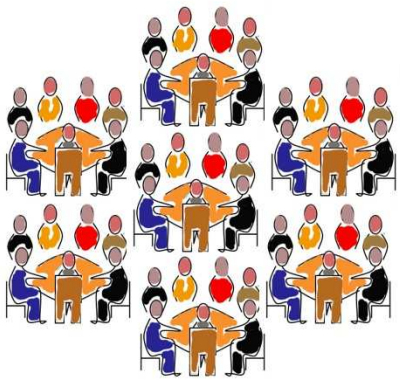

More glucose control
Less glucose control
Weight
Heart failure risk ratio (95% CI)
Events Total Events Total
PPAR agonists
2005 PROactive[36] 281 2605 198 2633 9·5% 1·43 (1·21–1·71)
2006 ADOPT[37] 22 1456 28 2895 3·6% 1·56 (0·90–2·72)
2006 DREAM[38] 14 2635 2 2634 0·7% 7·03 (1·60–30·90)
2009 BARI2D[41] 248 1183 218 1185 9·7% 1·14 (0·97–1·34)
2009 RECORD[42] 61 2220 29 2227 4·8% 2·10 (1·35–3·27)
2014 AleCardio[46] 122 3616 100 3610 7·7% 1·22 (0·94–1·59)
Subtotal 13715 15184 35·9% 1·42 (1·15–1·76)
Heterogeneity: Tau²=0·04; χ²=13·82, df=5; p=0·017; I²=64%
Test for overall effect: Z=3·29; p=0·0010
DPP-4 inhibitors
2013 EXAMINE[16,17] 106 2701 89 2679 7·3% 1·19 (0·90–1·58)
2013 SAVOR-TIMI 53[15] 289 8280 228 8212 9·5% 1·27 (1·07–1·51)
Subtotal 10981 10891 16·8% 1·25 (1·08–1·45)
Heterogeneity: Tau²=0·00; χ²=0·15, df=1; p=0·70; I²=0%
Test for overall effect: Z=2·94; p=0·0033
Intensive control
1998 UK Prospective Diabetes Study[35] 80 2729 36 1138 5·5% 0·91 (0·62–1·34)
2008 ACCORD[39] 152 5128 124 5123 8·3% 1·18 (0·93–1·49)
2008 ADVANCE[40] 220 5571 231 5569 9·3% 0·95 (0·79–1·14)
2009 VADT[43] 76 892 82 899 6·7% 0·91 (0·66–1·25)
Subtotal 14320 12729 29·8% 1·00 (0·88–1·13)
Heterogeneity: Tau²=0·00; χ²=2·80, df=3; p=0·42; I²=0%
Test for overall effect: Z=0·01; p=0·99
Insulin glargine
2012 ORIGIN[44] 310 6264 343 6273 9·8% 0·90 (0·77–1·05)
Subtotal 6264 6273 9·8% 0·90 (0·77–1·05)
Heterogeneity: not applicable
Test for overall effect: Z=1·34; p=0·18
Weight loss
2013 Look-AHEAD[45] 99 2570 119 2575 7·7% 0·80 (0·62–1·04)
Subtotal 2570 2575 7·7% 0·80 (0·62–1·04)
Heterogeneity: not applicable
Test for overall effect: Z=1·67; p=0·10
Total 47850 47652 100% 1·14 (1·01–1·30)
Heterogeneity: Tau²=0·04; χ²=45·56, df=13; p<0·0001; I²=71%
Test for overall effect: Z=2·04; p=0·041
Test for subgroup differences: χ²=21·85, df=4; p=0.00021, I²=81·7%
0·2
0·5
1
2
5
Favours glucose-lowering
Favours standard care

Exploring sources of heterogeneity (2)

- Meta-regression
  - Form of subgroup analysis that allows consideration of continuous variables, e.g. year of publication
  - Still allows more precise studies greater weight
  - Accounting for heterogeneity in terms of study/participant characteristics, but still likely to have unexplained heterogeneity – **random effects meta-regression**

Subgroup analyses & meta-regression

- Make sure enough studies ~ 10 for meta-regression
- Is there scientific rationale for each characteristics?
- Pre-specify small number of study/participant characteristics
  - Reduces chance of spurious findings
- If not pre-specified be clear analyses are *post hoc*
- Be aware of aggregation bias (ecological fallacy/bias)



Summary

- Understanding heterogeneity between studies is important
- Fixed or random effects meta-analysis = average effect
- Importance in interpretation and usefulness

- Methods for exploring heterogeneity are not ideal, but can be useful
- Hypothesis-generating rather than definitive