UMC Utrecht

Epidemiology and Big Data (INFOMEBD)

Outline of this presentation

- Principles of IPD meta-analysis
- Performing subgroup analysis

UMC Utrecht

Principles of IPD meta-analysis

Data structure IPD meta-analysis

Example of individual participant data from 10 hypertension trials that assess effect of treatment versus placebo on systolic blood pressure

Study ID	Patient ID	Age (years)	Sex (1=male, 0=female)	Treatment group (1=treatment, 0=control)	Systolic blood pressure before treatment (mm Hg)	Systolic blood pressure after treatment (mm Hg)
1	1	46	1	1	137	111
1	2	35	1	0	143	133
...	...	...	...	...	...	...
1	1520	62	0	0	209	219
2	1	55	0	1	170	155
2	2	38	1	1	144	139
...	...	...	...	...	...	...
2	368	44	1	0	153	129
3	1	51	1	1	186	166
3	2	39	0	1	201	144
...	...	...	...	...	...	...
3	671	54	0	0	166	141
...	...	...	...	...	...	...
10	1	71	0	1	149	128
10	2	59	1	0	168	169
...	...	...	...	...	...	...
10	978	63	0	1	174	128

Dotted line indicates where non-displayed rows of data occur.

Hypothetical data based on Wang et al.[27]

Key feature IPD

- IPD collected in different studies with likely differences in protocol, population, study features, interventions, outcomes

- Analysing the IPD as if from a single trial ignores these differences. Consequences:
  - differences in pooled estimates
  - differences in precision
  - no insight in heterogeneity of results between studies

Main approaches in IPD meta-analysis

- Take clustering of data into account through:
  - two stage approach
  - one stage approach

- Two stage: analyze each study separately and then combine results using traditional meta-analytical approaches

- One stage: analyze the individual data of all studies combined with a single model, but take clustering into account

2-stage IPD meta-analysis: Principles

- Stage 1: Estimate the effect estimate of interest and its precision (standard error) in each trial like odds ratio, relative risk, hazard ratio, mean difference
- Stage 2: calculate pooled (weighted) effect estimate across studies using traditional methods (fixed or random effects model)
- Statistical output looks similar to aggregate data meta-analysis
- Display individual results and summary measure in forest plot

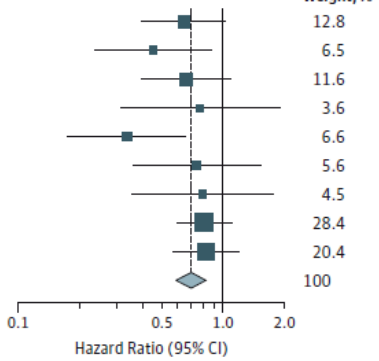
Simmonds et al. Meta-Analysis of individual patient data from Randomized Trials: A review of methods used in practice. Clinical Trials 2005:2;209-17.

2 stage IPD meta-analysis

Figure 2. Random Effects Meta-analyses Comparing Mindfulness-Based Cognitive Therapy (MBCT) With Other Variables

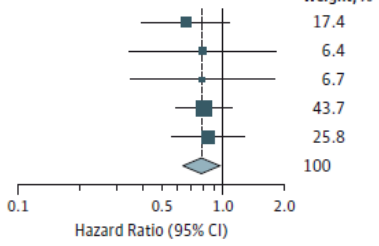
A MBCT vs no MBCT

Study	Hazard Ratio (95% CI)	Weight, %
Teasdale et al,[7] 2000	0.64 (0.40-1.03)	12.8
Ma and Teasdale,[6] 2004	0.45 (0.23-0.88)	6.5
Kuyken et al,[13] 2008	0.66 (0.40-1.08)	11.6
Bondolfi et al,[15] 2010	0.77 (0.31-1.90)	3.6
Godfrin and van Heeringen,[17] 2010	0.34 (0.17-0.66)	6.6
Segal et al,[18] 2010	0.74 (0.36-1.52)	5.6
Huijbers et al,[19] 2015	0.80 (0.36-1.78)	4.5
Kuyken et al,[21] 2015	0.81 (0.59-1.10)	28.4
Williams et al,[23] 2014	0.82 (0.57-1.20)	20.4
Overall (I^2 = 1.7%, P = .42)	0.69 (0.58-0.82)	100



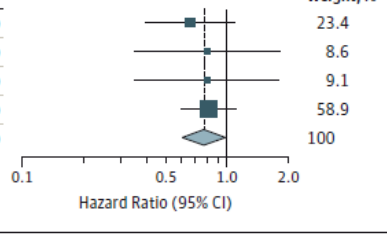
B MBCT vs any active treatment

Study	Hazard Ratio (95% CI)	Weight, %
Kuyken et al,[13] 2008	0.66 (0.40-1.08)	17.4
Segal et al,[18] 2010	0.80 (0.35-1.82)	6.4
Huijbers et al,[19] 2015	0.80 (0.36-1.78)	6.7
Kuyken et al,[21] 2015	0.81 (0.59-1.11)	43.7
Williams et al,[23] 2014	0.85 (0.56-1.28)	25.8
Overall (I^2 = 0.0%, P = .96)	0.79 (0.64-0.97)	100



C MBCT vs antidepressants

Study	Hazard Ratio (95% CI)	Weight, %
Kuyken et al,[13] 2008	0.66 (0.40-1.08)	23.4
Segal et al,[18] 2010	0.80 (0.35-1.82)	8.6
Huijbers et al,[19] 2015	0.80 (0.36-1.78)	9.1
Kuyken et al,[21] 2015	0.81 (0.59-1.11)	58.9
Overall (I^2 = 0.0%, P = .92)	0.77 (0.60-0.98)	100



Forest plot of 2-stage meta-analysis of aggregate data on hazard ratio scale comparing (A) risk of relapse of depression in participants receiving MBCT with participants not receiving MBCT; (B) risk of relapse of depression in participants receiving MBCT with participants receiving an alternative active therapy; and (C) risk of relapse of depression in participants receiving MBCT with participants receiving antidepressant medication. Weights are from random effects analyses.

2-stage IPD meta-analysis

- Benefits:
  - intuitive and straightforward approach
  - apply standard approaches for meta-analysis
  - similar measures, output & presentation
  - through forest plots: insight in heterogeneity
  - combination of IPD and AD technically straightforward
- Limitations:
  - potential for bias with few studies and/or few events per study
  - less flexible & power when examining patient- and study-level modifiers
  - survival data & prediction data

1-stage IPD meta-analysis: Principles

- Single regression model taking clustering within studies into account
- Various names: multilevel models, hierarchical models, mixed effects models
- Can explore simultaneously impact of trial and patient characteristics on treatment effect
- Requires greater statistical expertise
- Output will look different
- Results as in original studies:
  - RD, RR, OR, mean difference

1- vs. 2-stage IPD meta-analysis

- 2-stage more common, but 1-stage increasing
- 1- and 2-stage frequently give similar results
- 1-stage can provide greater flexibility and power when examining interactions and non-linear relationships with outcomes
- Sometimes 2-stage can be biased when having few and/or small studies
- 1-stage more complex and more choices during modelling
- 2-stage analysis a useful addition to 1-stage and vice versa

UMC Utrecht

IPD Review Benefits: Subgroup Analysis

Variation in treatment effect (single trial)

- Trials typically focus on the average effect
- Clinical intuition/experience: some patients benefit more than average
- Does the intervention effect vary by subgroups?
- Such variation is known as interaction by statisticians and as effect modification by epidemiologists

Variation in treatment effects (across studies)

- Many potential factors may cause variation
- Two main types of factors:
  - Patient factors: factors measured at patient level, values differ between patients
  - Study factors: factors measured at study level, same value for patients within studies, different between studies

- Patient factors: disease stage, presence of comorbid condition, age, gender
- Study factors: design features, region in the World, different dose or administration

AD meta-analysis: meta-regression

- The outcome variable is the effect estimate (like mean difference, log odds ratio or log risk ratio)
- The explanatory variables (covariates) are characteristics of studies that might influence the size of intervention effect

- Y = int $\longrightarrow$ overall pooling, 1 pooled estimate
- Y = int +b1*X1
  - b1=change in effect measure from one category to another or 1 unit change in continuous covariate
  - test for b1=0 is test for significance of covariate

AD meta-analysis: study-level covariate

- Study-level covariate:
  - Y = int ⟶ overall pooling, 1 pooled estimate
  - Y = int +b1*non-blinding
    (1=non-blinded, 0=blinded)

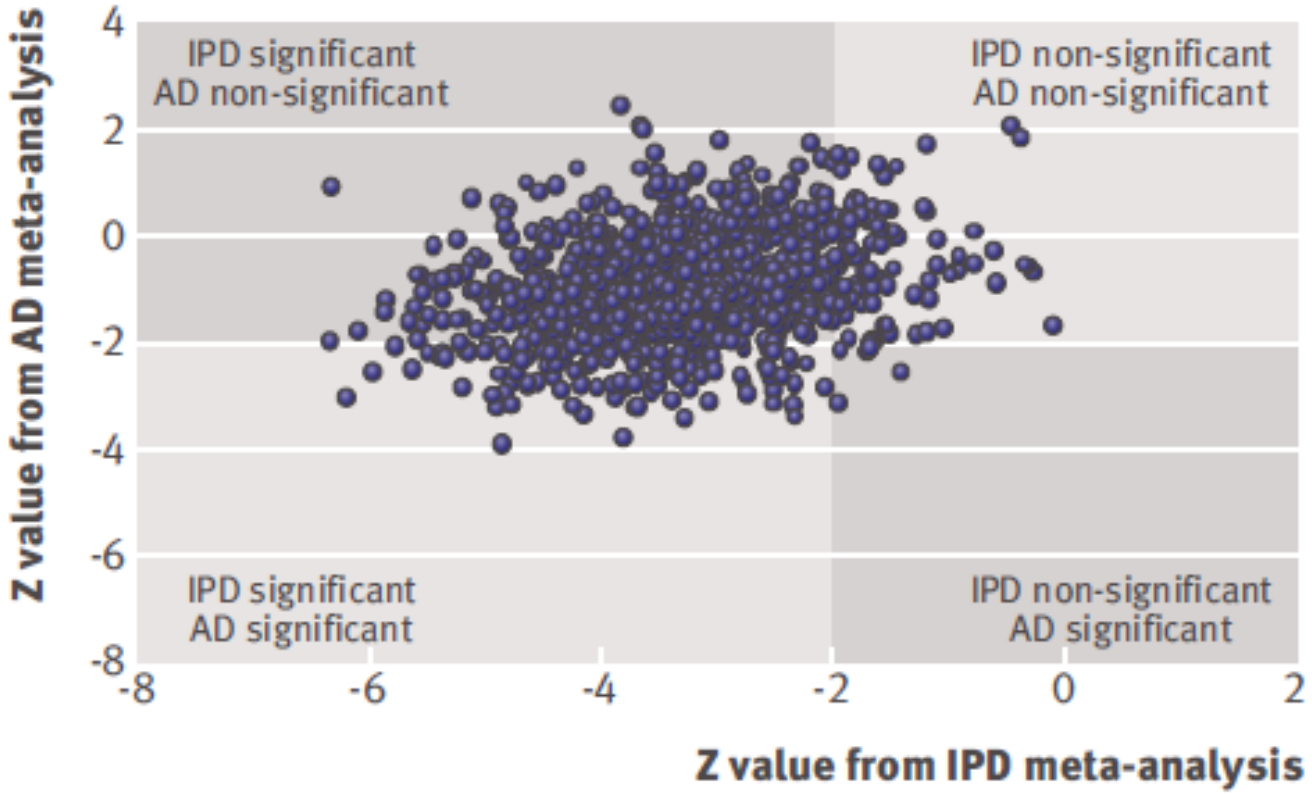
AD meta-analysis: summary measure of patient-level covariate

- Summary measure of patient-level covariate:
  - Y = int ⟶ overall pooling, 1 pooled estimate
  - Y = int +b1*%males

  Or
  - Y = int +b1*mean_age

- Investigate whether the effect measure (e.g. OR) varies with the % male or mean age across studies
- Indirect way: loose power & result may be invalid

IPD vs AD review: power subgroup analysis



Power:
IPD 91%
AD 15%

Fig 2 |Comparison of the power of meta-analyses to detect a differential treatment effect across two groups of patients when individual participant data (IPD) or aggregate data (AD) are used. Adapted from Lambert et al[19]

Ecological fallacy / Simpson's paradox

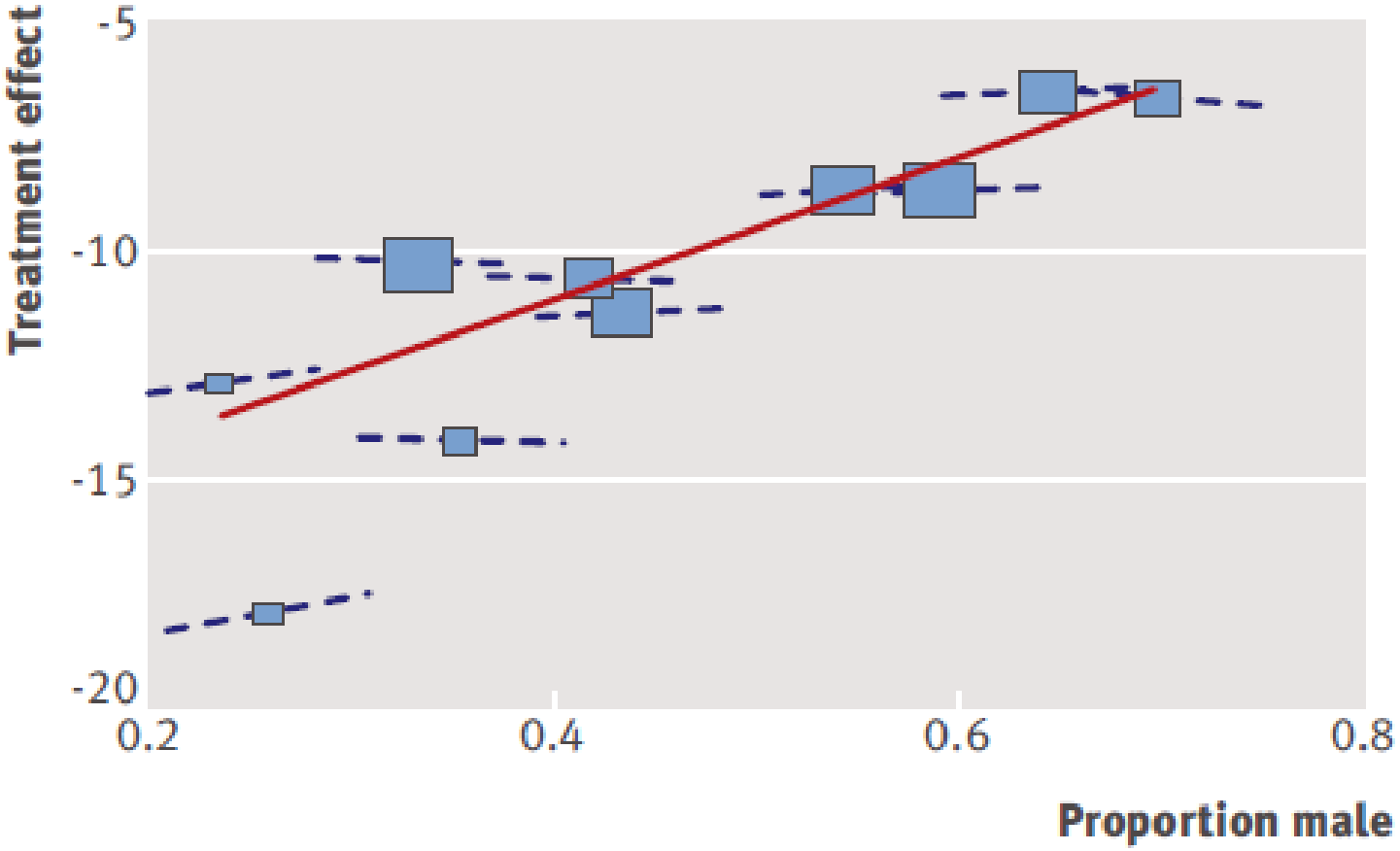


Fig 3 |An example of ecological bias within an aggregate data meta-analysis

Careful modelling of subgroups

RESEARCH METHODS AND REPORTING

Cite this as: *BMJ* 2017;356:j573
http://dx.doi.org/10.1136/bmj.j573

Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach?

David J Fisher,[1] James R Carpenter,[1,2] Tim P Morris,[1] Suzanne C Freeman,[1] Jayne F Tierney[1]

Research Article

Statistics in Medicine

Received: 14 April 2016, Accepted: 28 October 2016 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.7171

One-stage individual participant data meta-analysis models: estimation of treatment-covariate interactions must avoid ecological bias by separating out within-trial and across-trial information

Hairui Hua,[a] Danielle L. Burke,[b] Michael J. Crowther,[c,d] Joie Ensor,[b] Catrin Tudur Smith[e] and Richard D. Riley[b*†]

Benefits of subgroup analyses in IPD

- Analyze subgroups at the individual level across all studies (if measured)
- Greater power by analyzing subgroups at the level of the patient
- Flexible modelling of continuous covariates
- Analyze subgroups in a consistent way
- Examine combination of factors
- Possibility to adjust for imbalance in important prognostic factors between subgroups

UMC Utrecht

Summary

Summary

- Hierarchical nature (clustering) of data within a IPD meta-analysis needs to be taken into account
- Two-stage and one-stage approaches each having their pros and cons, but can also complement each other
- IPD meta-analysis offer greater flexibility, validity and power to examine subgroup effects (interactions between treatment and patient-level covariates). Still need sufficient data!
- Benefits of IPD even greater when developing and validating multivariable prediction models (more later this week)
- IPD no cure for poorly designed studies