Statistical methods for individual participant data meta-analysis

Catrin Tudur Smith[1] and Richard Riley[2]

[1]MRC North West Hub for Trials Methodology Research,
Department of Biostatistics,
University of Liverpool

[2] Unit of Public Health, Epidemiology & Biostatistics,
University of Birmingham

Acknowledgements

- Alex Sutton, University of Leicester
- Tony Marson, University of Liverpool
- Ruwanthi Kolamunnage-Dona, University of Liverpool
- Paula Williamson, University of Liverpool

Outline

- Introduction to IPD
- Statistical methods
  - Two-stage
  - One-stage
- Examining covariates
- Bias in meta-analysis

Not Covered in detail

- How to collect IPD
- How to manage IPD
- How to check IPD

See the following examples for further details:

Stewart, L. A., Tierney, J. F. and Clarke, M. (2008) Reviews of Individual Patient Data, in Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series (eds J. P. Higgins and S. Green), John Wiley & Sons, Ltd, Chichester, UK. doi: 10.1002/9780470712184.ch18

Stewart LA, Clarke MJ, on behalf of the Cochrane Working Party Group on Meta-analysis using Individual Patient Data. Practical methodology of meta-analyses (overviews) using updated individual patient data. Statistics in Medicine 1995;142057-79

Participant Experiences

- Are you undertaking an IPDMA?
  - Yes → why IPD?
  - No → reason for coming to the workshop?
- How successful have you been at getting IPD?
- What methods did you use for doing so?
- What sort of outcomes do you have in your review?
- What are your views of benefits / disadvantages of trying to do an IPD analysis?

Aggregate Data (AD) published

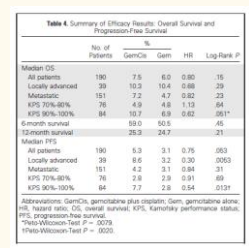


Table 4. Summary of Efficacy Results: Overall Survival and Progression-Free Survival

	No. of Patients	% GemCis	% Gem	HR	Log-Rank *P*
Median OS					
All patients	190	7.5	6.0	0.80	.15
Locally advanced	39	10.3	10.4	0.88	.29
Metastatic	151	7.2	4.7	0.82	.23
KPS 70%-80%	76	4.9	4.8	1.13	.64
KPS 90%-100%	84	10.7	6.9	0.62	.061*
6-month survival		59.0	50.5		.45
12-month survival		25.3	24.7		.21
Median PFS					
All patients	190	5.3	3.1	0.75	.053
Locally advanced	39	8.6	3.2	0.30	.0063
Metastatic	151	4.2	3.1	0.84	.31
KPS 70%-80%	76	2.8	2.9	0.91	.69
KPS 90%-100%	84	7.7	2.8	0.54	.0131

Abbreviations: GemCis, gemcitabine plus cisplatin; Gem, gemcitabine alone; HR, hazard ratio; OS, overall survival; KPS, Karnofsky performance status; PFS, progression-free survival.
*Peto-Wilcoxon Test $P = .0079$.
†Peto-Wilcoxon Test $P = .0020$.

Journal of clinical oncology 2006, *24:3946-3952.*

Individual participant data (IPD)

Patient Number	Treatment	Survival Time (Days)	Status	Age	Sex	Stage
1	E	44	Dead	67	m	IV
2	E	54	Dead	64	m	III
3	E	67	Alive	55	f	III
4	C	43	Dead	79	f	IV
5	C	70	Alive	62	m	IV
6	E	88	Dead	60	f	IV
7	C	99	Alive	57	m	III
8	C	45	Dead	66	m	III
9	E	90	Alive	59	f	III
10	C	23	Dead	53	m	IV

Aggregate Data (AD) requested

- Be aware there is another option
- Specific summary tables / statistics beyond those in paper may be sufficient / more desirable than relying on published result to carry-out the analysis required
- Desirable if investigators are unwilling to hand over their dataset but would supply further information

Why IPD?

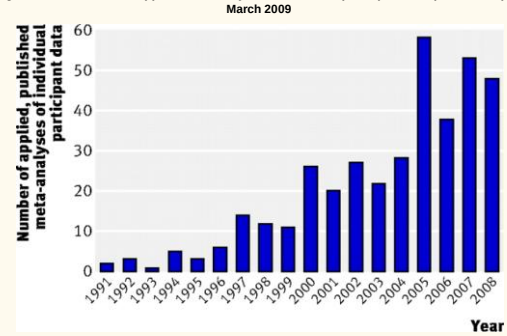
M-A of IPD is considered the gold standard approach to analysis

- Used relatively infrequently
- But, becoming more common (Simmonds et al 2005) – 79 IPD reviews on the IPDMWG site

General disadvantages

- Time consuming and costly
- May not be able to obtain all IPD - retrieval bias

Fig 1 Number of distinct, applied meta-analyses of individual participant data published up to March 2009



Riley R D et al. BMJ 2010;340:bmj.c221

BMJ

Why IPD?

M-A of IPD is considered the gold standard approach to analysis

- Used relatively infrequently
- But, becoming more common (Simmonds et al 2005) – 79 IPD reviews on the IPDMWG site

General disadvantages

- Time consuming and costly
- May not be able to obtain all IPD - retrieval bias

Why IPD?

Data checking & standardisation of analysis	• Outcome definition can be standardised across trials
More complete analysis	• Include follow-up beyond initial publication • Reinstate patients into the analysis who were originally excluded • Overcome outcome reporting bias
Detailed exploration of participant level covariates influence on treatment effect	• Maximum information using patient as unit of analysis - more power to identify clinically moderate interaction • Direct interpretation for individual patient • No reporting bias of subgroup analyses • No ecologic bias
More thorough analysis of time-to-event data	• Check model assumptions eg proportional hazards • More accurate

Examples of when IPD ?

- High patient exclusion rate
  - IPD meta-analysis in soft tissue sarcoma (1997), authors recovered data on 99% of the 344 patients that had been excluded from individual trial analyses.
  - Without additional patients HR=0.90 (p>0.05); including additional patients HR=0.85 (p<0.05)
- Time-to-event data
  - Epilepsy example, time to 12 month remission and time to treatment failure recognised as outcomes of clinical importance
  - Most trials reported different outcomes (50% reduction in seizures) or different definitions (time from specific dose level rather than time from randomisation), or
  - Trials did not report sufficient summary data to allow HR to be estimated reliably

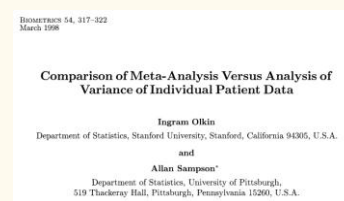
Examples of when IPD ?

- When interactions with treatment are important
  - Interaction between epilepsy type and treatment
  - Heterogeneity across studies in AD meta-analysis of cervical cerclage
- To investigate longer-term outcomes
  - such as for chronic diseases where events take place over a long period of time
- Meta-analysis of prognostic factors studies
  - Use a (small) consistent set of adjustment factors across studies
  - Use a consistent cutpoint across studies, or produce continuous marker results.
  - Assess the benefits of using combinations of markers

IPD vs AD

- IPD and AD meta-analysis can be equivalent
  - if data are equivalent
  - If treatment effect measure are equivalent

IPD vs AD



BIOMETRICS 54, 317-322
March 1998

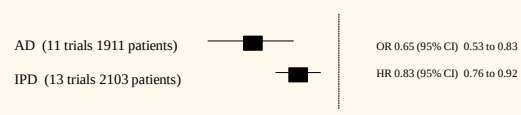
Comparison of Meta-Analysis Versus Analysis of Variance of Individual Patient Data

Ingram Olkin
Department of Statistics, Stanford University, Stanford, California 94305, U.S.A.
and
Allan Sampson*
Department of Statistics, University of Pittsburgh,
519 Thackeray Hall, Pittsburgh, Pennsylvania 15260, U.S.A.

IPD vs AD

- Discrepancies usually arise because IPD data sets include different data to AD
  - IPD may reinstate patients originally excluded
  - IPD may include additional follow-up data
  - IPD may use more appropriate effect measure

IPD vs AD



Pignon JP and Arriagada R. Lancet 1993.

IPD vs AD

"..the balance of gains and losses of the approach will vary according to the disease, treatment, and therapeutic questions explored" Stewart and Tierney 2002

- Empirical evidence - precision and size of effect varies compared to AD but no systematic pattern
- Further empirical evidence is needed :

  Individual patient data meta-analyses compared with meta-analyses based on aggregate data. Clarke MJ, Stewart L , Tierney J , Williamson PR Protocol for methodology review – Cochrane Library

SEE POSTER HERE!

Meta-Analysis of IPD

- Decisions for analysis with IPD
  - Two-stage or One-stage
  - Fixed or Random Treatment Effects
  - Which software?
  - Prognostic Factors and Effect Modifiers

Meta-Analysis of IPD

- Common approaches to M-A of IPD include
  - Two stage: create summary statistics out of IPD (stage 1) and combine using standard meta-analysis method (stage 2)
    - *either fixed effect or random effects approach*
  - One stage: combine all patient data from all studies in one single model taking into account the clustering of patients within study
    - *either fixed effect or random effects approach*

Common practice

- **Simmonds et al (2005)**, n=44, 1999-2001
- 65% with <=10 trials
- two-stage methods most common
- poor reporting

- **Pignon et al (2007)**, lung cancer, n=9, -2006
- two-stage methods most common

- **Kolamunnage-Dona (2008),** n=79 (62 with data on number of trials), IPDMWG
- median 10 trials, range 2-63
- two-stage methods most common

Two-stage : time to event data

STAGE 1: For each trial separately, reduce the IPD to Aggregate Data

(i) Fit a separate Cox proportional hazards model (IPD) to each trial
e.g proc phreg (by trial) in SAS

(ii) Obtain an estimate of log hazard ratio and its standard error for each trial

⇒ THIS IS AGGREGATE DATA

Two-stage : time-to-event data

STAGE 2: Combine the AD using standard meta-analysis methods

(i) Enter data into meta-analysis software e.g. Revman
(ii) Undertake meta-analysis using generic inverse variance method (either fixed effect or random effects)
- Calculate I-squared as usual

(iii) Could extend to adjust for covariates within trials
(iv) Extend to meta-regression if required

Two-stage : time to event data

- ALTERNATIVE: Stratified log-rank analysis (Early Breast Cancer Trialists' Collaborative Group)

STAGE 1: Undertake a log-rank analysis for each trial to obtain estimates of

The log-rank statistic U_j and its variance V_j

STAGE 2: Combine over all trials using

$$\hat{\beta} = \frac{\sum_{j=1}^{J} U_j}{\sum_{j=1}^{J} V_j} \quad \text{with} \quad SE(\hat{\beta}) = \sqrt{1/\sum_{j=1}^{J} V_j}$$

Two Stage: continuous or binary data

Stage 1: Fit a separate model for all patients in each trial and extract estimates of treatment effect and standard error

Continuous Data

Normal linear regression model – estimate of difference in means and standard error

Binary Data

Logistic regression model – estimate of log odds ratio and standard error

Stage 2: Pool across trials using standard meta-analysis methods

Software for Two Stage Approach

- Using Revman (free)

Stage 1:

Use a standard Stats Package to obtain estimates of treatment effect and SE eg SAS (proc phreg), R (coxph), STATA (stcox)

Stage 2:

Input data using Generic Inverse Variance Method in Revman

Note: 'O-E and Variance' option in Revman fits 'Peto' method if logrank 'O – E' and 'V' statistics have been obtained.

This is a FE analysis – no equivalent RE analysis is available in RevMan.

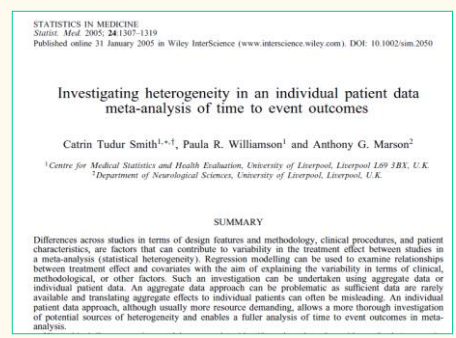
Software for Two Stage Approach

- Using SCHARP - Survival Curve and HAzard Ratio Program
- an interactive SAS-based application
- Analyses and plots IPD meta-analyses
- Uses two-stage approach
- Analysis of time-to-event, dichotomous and continuous outcomes with choice of measures
- Fixed effect and random effects models
- Version 4 available but still need SAS and SCHARP may still have bugs – limited support available
- Free from MRC Clinical Trials Unit, London

Two-stage Approach

- Benefits:
  - Straightforward
  - Accommodate Fixed and Random Effects (using DerSimonian-Laird method in usual way)
  - 'Standard' meta-analysis interpretation: Forest Plot and Heterogeneity statistics output
  - Can easily incorporate both IPD and AD estimates within the same meta-analysis
- Limitations
  - Cumbersome with many trials
  - Limited - cannot fully investigate patient level effect modifiers

One-stage regression models for time-to-event data



STATISTICS IN MEDICINE
Statist. Med. 2005; **24**:1307–1319
Published online 31 January 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2050

Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes

Catrin Tudur Smith[1,*,†], Paula R. Williamson[1] and Anthony G. Marson[2]

[1]*Centre for Medical Statistics and Health Evaluation, University of Liverpool, Liverpool L69 3BX, U.K.*
[2]*Department of Neurological Sciences, University of Liverpool, Liverpool, U.K.*

SUMMARY

Differences across studies in terms of design features and methodology, clinical procedures, and patient characteristics, are factors that can contribute to variability in the treatment effect between studies in a meta-analysis (statistical heterogeneity). Regression modelling can be used to examine relationships between treatment effect and covariates with the aim of explaining the variability in terms of clinical, methodological, or other factors. Such an investigation can be undertaken using aggregate data or individual patient data. An aggregate data approach can be problematic as sufficient data are rarely available and translating aggregate effects to individual patients can often be misleading. An individual patient data approach, although usually more resource demanding, allows a more thorough investigation of potential sources of heterogeneity and enables a fuller analysis of time to event outcomes in meta-analysis.

One-stage regression models for time-to-event data

Data from each patient in each trial are included in a single model eg Cox model

Fixed Effect (SFE/FE)

Cox model stratified by trial with fixed treatment effect

- Assume proportional hazards within trials, but not across trials
- Treatment heterogeneity assessed via trial-specific effects
- No allowance is made for residual heterogeneity

Random Effects (SFE/RE)

Cox model stratified by trial with random treatment effects

- Assume proportional hazards within trials, but not across trials
- Treatment heterogeneity assessed via heterogeneity parameter
- Allowance is made for residual heterogeneity

Other alternatives described by Tudur Smith et al 2005.

Software for one-stage regression models for time-to-event data

- Cox models with fixed effects fitted using standard statistical software eg proc phreg in SAS, coxph in R, stcox in STATA
- Random effects models
  - SAS IML: approach outlined by Yamaguchi (1999), adapted by Tudur Smith (2005) (fixed trial, stratified or random trial)
  - R coxme: still under development ?
- Abrams note that you can re-formulate Cox model as Poisson regression model (Whitehead, 1980;Lindsey, 1995)
  - Relatively easy to specify random effects
  - Implemented in R using lmer function
- Other estimation methods for random effects models - Cortinas Abrahantes et al (2007) compared 4 methods but were not able to make any clear recommendation

Comparison of methods for time-to-event data

- Comparison of five alternative 'one-stage' Cox models[1]
- Stratified Cox model appropriate for most situations
  - More appropriate to assume different baseline hazard in each trial rather than a common baseline across all trials
  - Trial effects as dummy variables compares patients across trials
  - Trial effects as random effects may not be reasonable?
  - Computationally more efficient to fit for random treatment effects
- NOTE: if many trials included may produce unstable estimates,
  - Efficiency gains for random trial effects greatest for moderate to large numbers of very small groups (of sizes two or three)

[1] *Tudur Smith and Williamson (Stat Med 2005)*

Comparison of methods for time-to-event data

- How does the stratified Cox model (one-stage) compare with two stage approaches ?

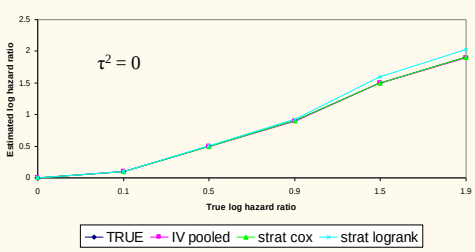
Clinical Trials
http://ctj.sagepub.com

A comparison of methods for fixed effects meta-analysis of individual patient data with time to event outcomes
Catrin Tudur Smith and Paula Ruth Williamson
Clin Trials 2007; 4; 621
DOI: 10.1177/1740774507085276

The online version of this article can be found at:
http://ctj.sagepub.com/cgi/content/abstract/4/6/621

- Stratified Cox versus Stratified logrank versus IV Cox
- All assuming fixed treatment effects
- Simulation study: 5 trials, 100 patients in each group,1000 simulations

Mean of log hazard ratios over 1000 simulations



Comparison of methods for time-to-event data

- How does the stratified Cox model (one-stage) compare with two stage approaches ?

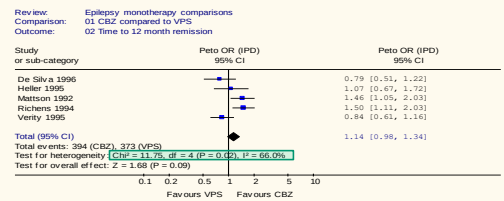
- Stratified Cox versus Stratified logrank versus IV Cox
- All assuming fixed treatment effects
- Simulation study: 5 trials, 100 patients in each group,1000 simulations

- No heterogeneity
  - all methods perform well for small effects as expected theoretically
  - stratified logrank displays bias and poor coverage for larger effects
- Increasing heterogeneity
  - Coverage decreases quite dramatically, bias increases

An example from epilepsy

- 5 RCTs with 1225 patients comparing CBZ and SV monotherapy (3 other eligible trials did not collect seizure data)
- Time to 12-month remission - not reported sufficiently in any of the trials
- Clinically important covariates: age, gender, epilepsy type, log(no. seizures), time from first ever seizure

Epilepsy: 12-month remission



Epilepsy: 12-month remission

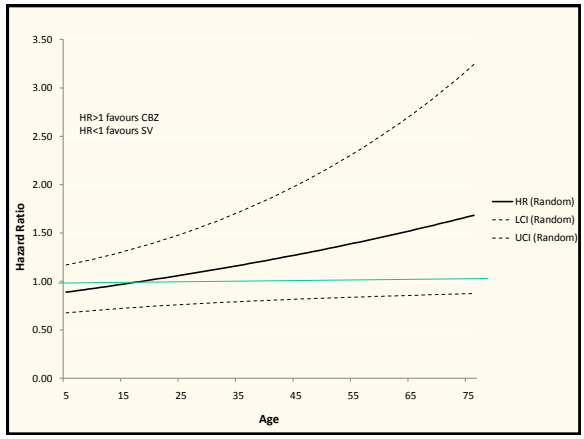
- Model without covariates

Heterogeneity: p=0.02, I^2 = 66%
SFE/FE: log HR = 0.132 (0.073)
SFE/RE: log HR = 0.098 (0.125), τ^2 = 0.0484 (0.055)

- Adjustment for: age, epilepsy type, log(seizures), age*treatment

SFE/FE: log HR = -0.162 (0.129)
SFE/RE: log HR = -0.163 (0.139), τ^2 = 0.006 (0.027)

- τ^2 reduced by 88%



One-stage models for other data

1. **Continuous Outcomes** (see Higgins JPT. et al. Stat Med 2001)
- Fixed effect - standard ANOVA model
- Random effects - SAS PROC MIXED, MLwiN, STATA xtmixed, winBUGS

2. **Binary Outcomes** (see Turner RM. et al. Stat Med 2000)
- Generally based on logistic regression models
- Fixed effect models - standard stats software eg SAS, R, STATA
- Random effect models – MLwiN, STATA gllamm, winBUGS

3. **Ordinal Outcomes** (see Whitehead A. et al. Stat Med 2001)
- Based on proportional odds models

Summary

Decisions for analysis include:

- Two-stage or One-stage?
  - Two-stage may be more straightforward for simple cases
  - Useful if IPD not available for all trials
  - But limited: One-stage more flexible
- Fixed or Random Treatment Effects?
  - Usual consideration for meta-analysis
  - Accommodated by 1-stage or 2-stage
  - Random effects can be more complex in 1-stage framework
- Which software?
  - Depends on data and model
  - Random effects survival analysis using SAS IML but not very efficient
  - Multilevel model software for other outcomes
- Prognostic Factors and Effect Modifiers
  - Use one-stage approach....

Using IPD meta-analysis to examine treatment-covariate interactions

Richard D. Riley

Dept of Epidemiology, Public Health & Biostatistics, & School of Mathematics

University of Birmingham

e-mail: r.d.riley@bham.ac.uk

Many thanks to:
Paul Lambert

Outline

Aim:
to show a range of examples where IPD meta-analysis helps examine treatment-covariate interactions

- One-step and two-step approaches
- Differences to a meta-regression of summary data
- Understanding threat of ecological bias

Part 1:

Rationale and Methods

Stratified medicine

- Increasing interest in *personalised* or *stratified* medicine
- We want to tailor treatment to individuals, or to groups of similar individuals
- To do this, we need to identify individual-level factors (covariates) that *modify* treatment response
- Essentially, what factors cause some patients to respond better to treatment than others?

Stratified medicine

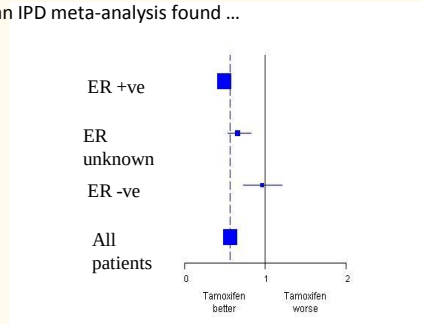
- For commissioners of healthcare
  - stratified medicine offers the potential to *maximise treatment related benefit* and *reduce treatment related harm.*
- For developers of new interventions
  - stratification may offer the opportunity to rescue a treatment which fails to show overall benefit in unselected patients, but that might have *worthwhile benefit in an identifiable subgroup*.

Stratified medicine

- Statistically, this means we want to examine & estimate so-called treatment-covariate interactions
  - *i.e. quantify how particular covariates interact with treatment effect*
- Also known as subgroup effects & effect modifiers
- Individual studies usually have *low power* to detect them, as they are powered on the overall treatment effect (the average across all individuals)
- By combining studies, meta-analysis thus offers an opportunity to increase power to detect true treatment-covariate interactions

Example: Estrogen receptor in breast cancer

• Tamoxifen is only given to patients who are ER positive, as an IPD meta-analysis found …



IPD Methods: A two-stage approach

• So how do we estimate treatment-covariate interactions in an IPD meta-analysis?

• Let us consider an example with continuous data

• There are 10 trials in hypertension (high blood pressure) and we are interested in the treatment effect on systolic blood pressure (SBP)

• Also interested in treatment-covariate interactions

Continuous data IPD

Study	Patient	SBP initial	SBP final	treat	placebo	age	sex
1	1	190	185	1	0	58	1
1	2	175	172	1	0	69	1
1	3	184	185	0	1	39	0
1	4	192	182	0	1	45	1
2	1	201	199	1	0	51	0
2	2	169	154	1	0	42	1
2	3	171	170	0	1	50	1
2	4	179	168	0	1	67	0
3	1	197	167	1	0	83	1
3	2	189	171	1	0	78	0
3	3	184	188	0	1	55	1
3	4	168	161	0	1	61	0

Can see that treatment effects, baseline factors and prognostic variables are available *per individual*

(note data are truncated for each study, as actually hundreds of patients)

IPD Methods: A two-stage approach

(a) Overall treatment effect

Step 1: Estimate the treatment effect and its variance in each IPD study using an appropriate method, such as analysis of covariance

Step 2: Take the effect estimates for each study, and combine them in a usual random-effects meta-analysis

- Gives a pooled treatment effect (across all individuals) of -10.16 (95% CI: -12.27 to -8.06).
- So hypertension treatment is significantly effective in reducing systolic blood pressure by, on average, 10.16 mm Hg more than control.

IPD Methods: A two-stage approach

(b) Effect of sex on the treatment effect

- Let 1 = males, and 0 = females

Step 1: Estimate the interaction between sex covariate and treatment effect, and its variance, in each IPD study separately using an appropriate method, such as analysis of covariance

Step 2: Take the interaction estimates for each study, and combine them in a usual fixed-effect of random-effects meta-analysis

IPD Methods: A two-stage approach

(b) Effect of sex on the treatment effect

- *sex* = 1 for males, and 0 for females;
- *treat* = 1 for treatment, and 0 for control
- Let i = study, and j = patient

STEP 1: $$SBP_{ij} = \phi_i + \beta_{1_i} SBP_{0ij} + \beta_{2i} sex_{ij} + \theta_i treat_{ij} + \gamma_i (sex_{ij} \times treat_{ij}) + \varepsilon_{ij} \qquad \varepsilon_{ij} \sim N(0, \sigma_i^2)$$

Treatment effect for females

Control effect — Treatment-sex interaction — Residual error

$\gamma_i =$ Change in treatment effect for males compared to females

IPD Methods: A two-stage approach

(b) Effect of sex on the treatment effect

- *sex* = 1 for males, and 0 for females;
- *treat* = 1 for treatment, and 0 for control
- Let i = study, and j = patient

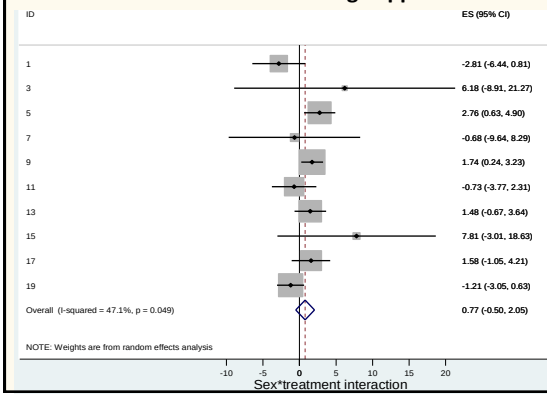
STEP 2:

e.g. fixed-effect meta-analysis

$$\hat{\gamma}_i = \gamma + \varepsilon_i \qquad \varepsilon_i \sim N(0, \text{var}(\hat{\gamma}_i))$$

Best estimate, across all studies, of the difference in treatment effect for males compared to females

IPD Methods: A two-stage approach



IPD Methods: A one-stage approach

- Alternatively, one can undertake a one-stage approach
- The IPD from all trials are analysed simultaneously
- Clustering of patients within trials accounted for
- Quicker and obtain multiple summary estimates together
- Obtains very similar estimates to two-stage approach

However …

- Including interactions requires *careful separation* of within-study and across-study relationships (Riley et al.)
- Essentially, you can explain both within-study variability and between-study variability … so need to separate these things out to avoid *ecological bias* (more later)

Part 2:

Examples of why the IPD approach is better than the aggregate data approach

If IPD are not available …

- Hope that study authors report the treatment-covariate interactions …. unfortunately rare
- Often one can only do a meta-regression

i.e. regress the study (i) treatment effect estimates (ϑ_i) against average patient-level covariates

$$\hat{\theta}_i = \alpha_i + \gamma(\text{proportionmale}) + u_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \text{V}(\hat{\theta}_i))$$

If IPD are not available …

- Hope that study authors report the treatment-covariate interactions …. unfortunately rare
- Often one can only do a meta-regression

i.e. regress the study (i) treatment effect estimates (ϑ_i) against average patient-level covariates

$$\hat{\theta}_i = \alpha_i + \gamma(\text{proportionmale}) + u_i + \varepsilon_i \qquad \varepsilon_i \sim N(0, \text{V}(\hat{\theta}_i))$$

Called the 'across-study interaction'. Tells us how much the average treatment effect differs in a study with only males compared to a study with only females

- crucially, this is different to the 'within-study interaction' obtained by analysing the IPD (Riley et al., 2008)

Within-study versus between-study interactions

Within-study interaction (from IPD)

- Effect of individual covariates on treatment effectiveness
- Results tailored to individual patient
- e.g. the treatment effect for males compared to females is …
- Explains within-study variability (residual error)

Across-study interaction

- How mean patient-level covariate in a study is associated with the mean treatment effect
- Results relate to the study-level (population)
- e.g. In a population with a proportion of 70% males, the underlying mean treatment effect is …
- Explains between-study variability

61

Within-study versus between-study interactions

- Within-study effects meaningful to individual patient
- But usually not obtainable if IPD not available
- Across-study effects meaningful at the population level
- Available when mean covariate is available for each study
- Simulation studies show that in *ideal* conditions across-study interactions will reflect within-study interactions ('unbiased')
- But across-study effects have low power, & prone to ecological bias & confounding across studies: Interpret with caution!
- e.g. studies with high proportion male may also have a higher dose of treatment; thus trend in treatment effect due to dose of drug and not proportion male

62

Example 1: Application to hypertension data

- How does being male modify treatment effect on SBP?
- Within-study effect
- γ_w = 0.77 (-0.5 to 2.05)

if for females the treatment reduces SBP by 20 mmHg more than placebo

then for males the treatment reduces SBP by 19.23 mmHg more than placebo

non-significant

63

Example 1: Application to hypertension data

- How does being male modify treatment effect on SBP?
- Within-study effect — Across-study effect
- γ_w = 0.77 (-0.5 to 2.05) — γ_A = 15.02 (8.98 to 21.1)

if for females the treatment reduces SBP by 20 mmHg more than placebo

then for males the treatment reduces SBP by 19.23 mmHg more than placebo

non-significant

if female studies have an underlying treatment effect that reduces SBP by 20 mmHg

then male studies have an underlying treatment effect that reduces SBP by 4.98 mmHg

significant

64

VERY DIFFERENT CONCLUSIONS, DUE TO ECOLOGICAL BIAS / CONFOUNDING

Example 1: Application to hypertension data

- How does being male modify treatment effect on SBP?
- Within-study effect — Across-study effect

?

NO IPD

γ_A = 15.02 (8.98 to 21.1)

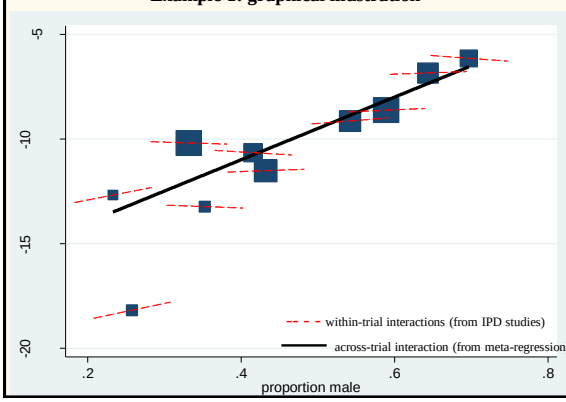
if female studies have an underlying treatment effect that reduces SBP by 20 mmHg

then male studies have an underlying treatment effect that reduces SBP by 4.98 mmHg

significant

65

Example 1: graphical illustration



Example 2: Increased power to detect true covariate interactions (Lambert et al., 2004)

• 1000 meta-analyses simulated, each with 5 trials and treatment effective for high risk patients but ineffective for low risk patients.

• Each meta-analysis analysed first using IPD, and then using meta-regression; treatment-covariate interactions estimated in both cases

• The % of 1000 meta-analyses that detect this true treatment-covariate interaction with statistical significance gives the power

• The % is usually far higher when using within-study interactions from IPD than when using interactions from meta-regression

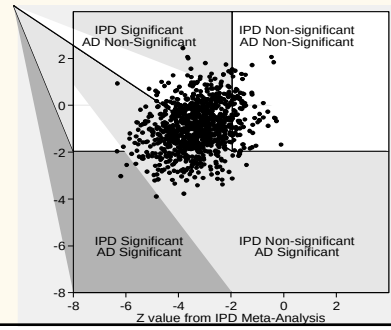
• Only when there is large between-study variation in the mean covariate value does the power of meta-regression appear adequate … *but even then ecological bias and confounding may occur*

67

Example 2: graphical illustration

IPD approach has a power of 90.8%

Meta-regression approach has a power of 10.8%



68

Example 3: Effect of elevated panel reactive antibodies on the effectiveness of anti-lymphocyte antibody induction (Berlin et al., 2022)

• Meta-analysis of five randomised trials of anti-lymphocyte antibody induction therapy for renal transplant patients

• Interested in the difference in treatment effect between patients with elevated antibodies compared to non-elevated

• A meta-regression is used to examine the across-trials interaction:

estimated difference in log odds of treatment failure between a trial with only elevated patients compared to a trial with only non-elevated patients = -0.01 (p = 0.68)

• Did the authors need IPD to obtain this result?

• What would you conclude from this about whether treatment effect is different for elevated and non-elevated patients?

Example 3: Effect of elevated panel reactive antibodies on the effectiveness of anti-lymphocyte antibody induction (Berlin et al., 2022)

• The reviewers also estimate the pooled within-study interaction

estimated difference between elevated and non-elevated patients in the log-odds of treatment failure = - 1.33 (p = 0.01)

• Did the authors need IPD to obtain this result?

• Suggest potential reasons why there is a substantial difference between within-study & across-study interactions

• Is there a genuine difference in treatment effect between elevated and non-elevated patients?

THANK YOU! ***r.d.riley@bham.ac.uk***

Useful References

• Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. BMJ 2010; 340: c221

• Higgins JP, Whitehead A, Turner RM, Omar RZ, Thompson SG: Meta-analysis of continuous outcome data from individual patients. Stat Med 2001; 20: 2219-4

• Berlin JA, Santanna J, Schmid CH, Szczech LA, Feldman HI. Individual patient- versus group-level data meta-regressions for the investigation of treatment effect modifiers: ecological bias rears its ugly head. *Stat Med* 2002;21(3):371-87

• Lambert PC, Sutton AJ, Abrams KR, Jones DR. A comparison of summary patient-level covariates in meta-regression with individual patient data meta-analysis. *J Clin Epidemiol* 2002;55(1):86-94.

• Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG: A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2000, **19**: 3417-3432.

• Tudur-Smith C, Williamson PR, Marson AG: Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med* 2005, **24**: 1307-1319.

• Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. J Research Synthesis Methods 2010

• Riley RD, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Stat Med* 2008; **27**: 1870-93

• Jones AP, Riley RD, Williamson PR, Whitehead A. Meta-analysis of individual patient data versus aggregate data from longitudinal clinical trials. *Clin Trials* 2009;6(1):16-27.

71

Are IPD meta-analyses biased?

Richard D. Riley

Dept of Epidemiology, Public Health & Biostatistics, & School of Mathematics

University of Birmingham

e-mail: r.d.riley@bham.ac.uk

Many thanks to:
Alex Sutton and Ikhlaaq Ahmed

72

Outline

Aim:

to discuss potential biases that may affect IPD meta-analyses and to provide examples

- Publication bias
- Selection bias
- Unavailable data

73

Part 1:

Possible biases

74

Is IPD meta-analysis really the gold-standard?

- The phrase 'gold-standard' is often used in articles to describe the IPD meta-analysis approach
- We have already discussed many reasons why IPD is preferable over a traditional meta-analysis of aggregate data from publications
- Yet there has been little consideration of how potential biases may impact upon IPD meta-analyses
- For example, biases may act in:
  - the identification of relevant studies,
  - the decision about which studies to seek IPD from,
  - the amount of IPD obtained from studies,
  - the type of studies that agree to provide their IPD

75

Publication & related biases

- ***Publication bias*** occurs when studies with statistically significant or clinically favourable results are more likely to be published than studies with non-significant or unfavourable results.
- Other related biases exist such as
  - time-lag bias
  - selective outcome reporting
  - language bias
  - duplication bias, etc
- Leads to meta-analyses which
  - synthesise an *incomplete* set of the evidence
  - produce summary results potentially *biased* toward favourable treatment effects.

76

Publication & related biases

- IPD allows trial results to be derived directly and independent to study reporting
  - it thus has potential to reduce publication and related biases
  - especially if IPD are obtained for unpublished trials
- Yet, all these bias problems hide pertinent trials and their results
- Thus – just as in a standard systematic review and meta-analysis – they may cause IPD researchers to miss relevant but non-significant trials
- **Hence, IPD from relevant trials may not be sought**

77

Publication bias: evidence for concern

- Burdett et al. found that in 9 of 11 IPD reviews, the meta-analysis result was *closer to the null* when 'gray literature' studies were included (e.g. unpublished trials, conference abstracts, etc)
- Ahmed et al. (submitted) examined 31 IPD meta-analyses of trials published between 2007 and 2009
  - Only 9 of the 31 articles included 'gray literature' IPD in their primary meta-analysis.
  - Thus majority (65%) do not include IPD from 'gray literature'
- This emphasises why obtaining IPD does not automatically remove the potential for publication related biases in meta-analysis.
- *Despite this, only 10 of the 31 articles discussed or examined statistically the threat of publication bias in their IPD meta-analysis*

78

Data availability bias

- IPD may not be available for all studies
- If unavailability of IPD is related to the study results, this may cause bias
- ***The impact of availability bias is hard to predict***

e.g. (1):
Researchers of studies with non-significant results may be more likely to have destroyed or lost their IPD
- *bias IPD meta-analyses toward a favourable treatment effect*

e.g. (2)
Researchers of studies with favourable findings may not provide their IPD because they want to utilise it further, for subgroup effects or an extended follow-up
- *bias IPD meta-analyses toward a lower treatment effect*

79

Data availability bias: evidence for concern

(i) Review of 199 applied IPD meta-analyses (Riley et al., 2006):

- 102 (58%) obtained IPD for > 90% of the studies
  - encouraging; IPD approach feasible
- 51 (29%) obtained IPD for < 80% of the studies
  - concerning; substantial evidence ignored

(ii) Review of 30 IPD meta-analyses of trials between 2007 and 2009 (Ahmed et al., submitted)

- 16 (53%) did not obtain all the IPD they asked for
- 10 (33%) obtained IPD from less than 80% of trials

Reasons include trial data being lost or destroyed; & study authors not being contactable, unwilling to collaborate, or unable to send their data

80

Data availability bias: evidence for concern

- Reasons for unavailable IPD include:
  - trial data being lost or destroyed;
  - study authors not being contactable, unwilling to collaborate or unable to send their data
- *When IPD are unavailable for some trials, the IPD approach may not be better than a meta-analysis of aggregated data from <u>all</u> trials.*
- Investigate the potential impact of non-IPD trials on IPD meta-analysis conclusions, wherever possible.

e.g. Vale et al. obtain aggregate results for three of their ten missing trials, and '*incorporating them into the meta-analysis did not materially change the results*'.

- Statistical approaches which synthesise both IPD and aggregate data are potentially valuable here (Riley et al., 2008)

81

Selection bias

- *Selection bias* occurs if reviewers only seek IPD from a subset of existing studies, and this subset does not reflect the evidence-base.
- This is a particular concern when:
  - relevant studies are not identified by a systematic review but rather through contacts or friends in the field
  - when selection takes place with *knowledge of study results*
- The impact of selection bias is hard to predict
  - it may (directly or indirectly) be affected by the selectors' knowledge of the field, their research contacts & collaborations, and their opinion about the research question of interest.
- It is less of a concern for prospective IPD meta-analysis: as study results are unknown at the time of study recruitment

82

Selection bias: evidence for concern

In the survey by Ahmed et al. (submitted)

- 22 of the 31 articles performed a systematic review to identify all relevant trials, from which IPD was then requested.
- In the other 9 articles selection bias is a potential concern, as identification of relevant studies was either not stated or based on a more selective, non-systematic approach.

For example, Papakostas et al. state:

'although we included all eligible studies sponsored by GlaxoSmithKline regardless of whether they have been published or not, it is possible that studies sponsored by other sources have been conducted but have not been yet published or presented at major scientific meetings.'

- *Do not automatically view an IPD meta-analysis as 'gold standard' without due thought as to how IPD studies were chosen.*

83

Part 2:

Illustrated examples

84

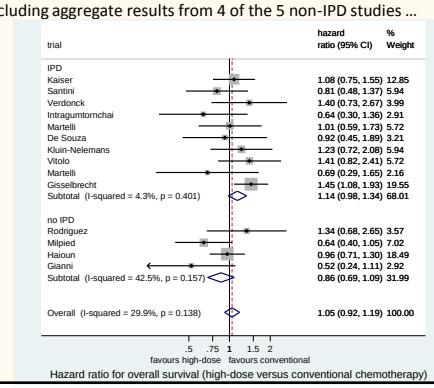
Example 1: Greb et al.

- Review whether high-dose chemotherapy with autologous stem cell transplantation as part of first-line treatment improves survival in adults with aggressive non-Hodgkin lymphoma.
- 15 randomised trials comparing high-dose versus conventional chemotherapy were identified by a systematic review.
- IPD were sought from all 15 trials, *so selection bias is not a concern.*
- However, publication and availability biases are a threat, as
  - all trials were fully published
  - IPD was unavailable for five of them (33%).
- Greb et al. examine both these issues; now extend their work …

85

Greb et al: Data availability bias?

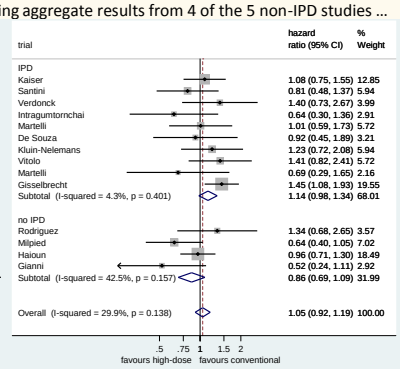
- Impact of including aggregate results from 4 of the 5 non-IPD studies …



Hazard ratio for overall survival (high-dose versus conventional chemotherapy)

Greb et al: Data availability bias?

- Impact of including aggregate results from 4 of the 5 non-IPD studies …



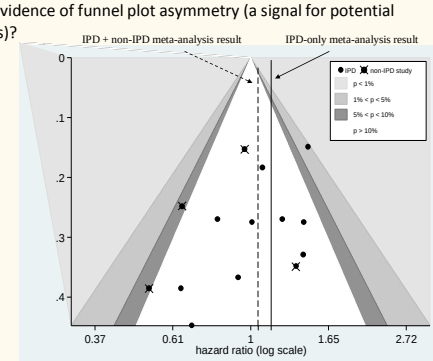
Non-IPD studies are mostly negative;

including them diminishes the effect of high-dose chemotherapy

Hazard ratio for overall survival (high-dose versus conventional chemotherapy)

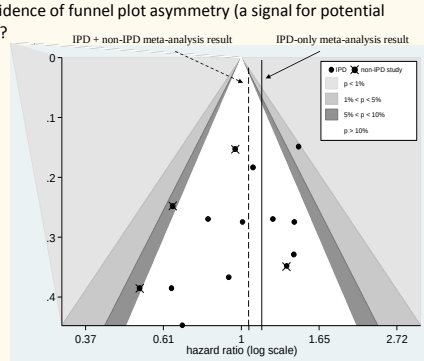
Greb et al: Publication bias?

- Is there any evidence of funnel plot asymmetry (a signal for potential publication bias)?



Greb et al: Publication bias?

- Is there any evidence of funnel plot asymmetry (a signal for potential publication bias)?



No statistical evidence of asymmetry either when including or excluding the non-IPD studies (p> 0.1)

- Any slight asymmetry is in bottom right, where significant studies would lie

- thus publication bias is not a concern

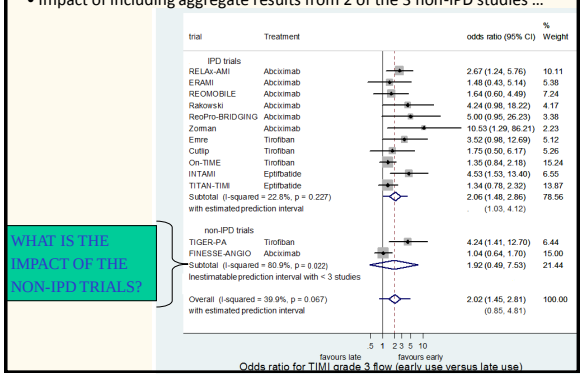
Example 2: De Luca et al.

- Review the benefits of early versus late use of Gp IIb-IIIa inhibitors in patients undergoing primary angioplasty for ST-segment elevation myocardial infarction.
- A primary angiographic endpoint was whether patients achieved a preprocedural Thrombolysis in Myocardial Infarction Study (TIMI) grade 3 flow distal embolisation.
- A systematic review identified 14 relevant trials and IPD was sought from them all, so selection bias is not a concern.
- However, availability and publication biases are a threat, as
  - IPD was unavailable for 3 trials (21%)
  - all 11 trials providing IPD were fully published.
- De Luca et al. did not investigate these biases, so let's do it …

90

De Luca et al: Data availability bias?

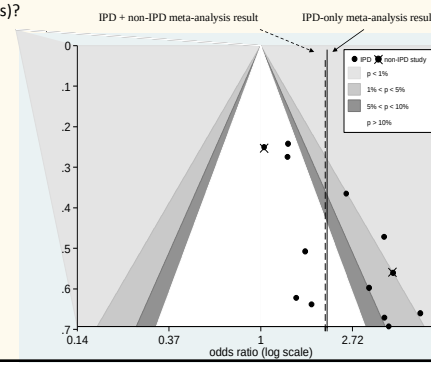
• Impact of including aggregate results from 2 of the 3 non-IPD studies …



WHAT IS THE IMPACT OF THE NON-IPD TRIALS?

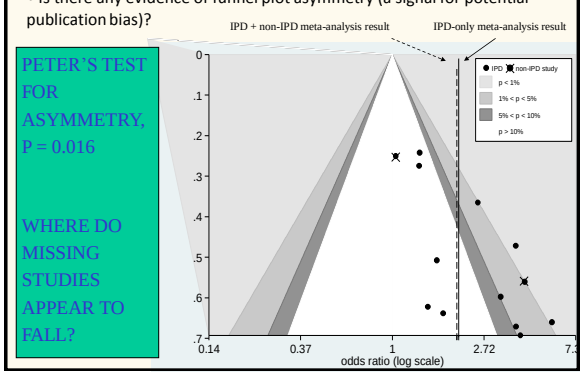
De Luca et al: Publication bias?

• Is there any evidence of funnel plot asymmetry (a signal for potential publication bias)?



De Luca et al: Publication bias?

• Is there any evidence of funnel plot asymmetry (a signal for potential publication bias)?



PETER'S TEST FOR ASYMMETRY, P = 0.016

WHERE DO MISSING STUDIES APPEAR TO FALL?

Many of the issues discussed today are in …

Cite this as: BMJ 2010;340:c221

Research Methods & Reporting

Meta-analysis of individual participant data: rationale, conduct, and reporting

Richard D Riley, *senior lecturer in medical statistics*[1], **Paul C Lambert**, *senior lecturer in medical statistics*[2], **Ghada Abo-Zaid**, *postgraduate student*[3]

[1] Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham B15 2TT, [2] Centre for Biostatistics and Genetic Epidemiology, Department of Health Sciences, University of Leicester, Leicester LE1 7RH, [3] School of Mathematical Sciences, University of Birmingham, Birmingham B15 2TT

Correspondence to: R D Riley r.d.riley{at}bham.ac.uk

The use of individual participant data instead of aggregate data in meta-analyses has many potential advantages, both statistically and clinically. **Richard D Riley and colleagues** describe the rationale for an individual participant data meta-analysis and outline how to conduct this type of study

94

Reporting IPD meta-analyses

Box 4: Suggested information to report from an IPD meta-analysis, to supplement those reporting guidelines elsewhere [34 35]

Articles of applied IPD meta-analyses should report:

- if there was a protocol for the IPD project, and where it can be found
- if ethics approval was necessary and (if appropriate) granted
- why the IPD approach was initiated originally
- the process used to identify relevant studies for the IPD meta-analysis
- how authors of relevant studies were approached for IPD
- how many studies (or collaborating groups) were approached for IPD, and the proportion that provided IPD
- whether those studies that provided IPD gave all their IPD or only a proportion; if the latter, then describe what information was omitted and why
- whether the IPD results for each study were comparable to published results, and if not why not (e.g. IPD contained updated or modified information)
- the number of patients within each of the IPD studies and, if appropriate, the number of events

95

Reporting IPD meta-analyses

- the number of studies that did not provide IPD, the reasons why, and the number of patients (and events) therein
- details of any missing individual-level data within the available IPD for each study, and how this was handled within the meta-analyses performed
- details and reasons for including (or excluding) patients into the analysis that were originally excluded (or included) by the study investigators
- whether a one-step or a two-step IPD meta-analysis was performed, and the statistical details thereof, including how clustering of patients within-studies was accounted for.
- how many patients were used from each study in each meta-analysis performed
- whether the assumptions of the statistical models were validated (e.g. proportional hazards) within each study
- any qualitative or quantitative differences between those studies providing IPD and those studies not providing IPD (if appropriate)
- the robustness of meta-analysis results to the inclusion/exclusion of non-IPD studies (if appropriate)
- how IPD and non-IPD studies were meta-analysed together (if appropriate)

96

THANK YOU! ***r.d.riley@bham.ac.uk***

Useful References

- Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: conduct, rationale and reporting. BMJ 2010; 340: c221
- Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Bmj* 2011;342:d4002
- Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof* 2002;25(1):76-97.
- Riley RD. Commentary: Like it and lump it? Meta-analysis using individual participant data. *Int J Epidemiol* 2010;39((5)):1359-1361.
- Stewart L, Tierney J, Burdett S. Do Systematic Reviews Based on Individual Patient Data Offer a Means of Circumventing Biases Associated with Trial Publications? In: Rothstein HR, Sutton AJ, Borenstein M, editors. *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. Chichester, UK.: John Wiley & Sons, Ltd., 2006.
- Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. *International Journal of Technology Assessment in Health Care* 2003;19(1):129-34
- Clarke MJ, Stewart LA. Obtaining data from randomised controlled trials: how much do we need for reliable and informative meta-analyses? In: Chalmers I, Altman DG, editors. *Systematic reviews*. London: BMJ Publishing, 1995:37-47
- Moreno SG, Sutton AJ, Ades AE, Stanley TD, Abrams KR, Peters JL, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Methodol* 2009;9:2